Quantifying Usability in Secure Graphics: Assessing the User Costs of Protecting 3D Content

Jiajun Zhu* Jonathan Z. Bakdash† David Koller* Thomas Banton† Dennis R. Proffitt† Greg Humphreys*

University of Virginia

Abstract

There is an increasing need for methods for secure dissemination of interactive 3D graphics content, providing protection for valuable 3D models while still allowing them to be widely shared. Existing systems for protected sharing of 3D models may introduce perturbations into the rendered images of the content, in order to defend against potential malicious reconstruction attacks that could otherwise recover the 3D model shape. However, the particular nature and magnitude of these perturbation defenses has not been based upon any rigorous analysis or measurement of their perceptual effect on non-malicious users of the protected graphics system. In this paper, we take the first steps toward such an analysis, conducting a series of user studies that evaluate the impact (as measured by user reaction time) of varying amounts of noise applied to user interactions in a real-time 3D rendering system. We are thus able to experimentally determine the most appropriate tradeoffs between noise perturbation defenses and the security of the 3D content against typical reconstruction attacks.

CR Categories: I.3.m [Computer Graphics]: Miscellaneous—Perception and Security

Keywords: secure graphics, perception, usability, 3D reconstruction

1 Introduction

As an increasing amount of interactive 3D graphics content is created, there is a growing need for tools that allow protected dissemination of this content. Creators of highly valuable digital 3D models may be reluctant to share their intellectual property widely unless methods exist to prevent theft or misuse of the 3D content. Notable examples include high-resolution digital 3D scans of prominent cultural heritage artifacts, such as famous statues or archaeological artifacts [Koller and Levoy 2005]. Other examples include trademarked character models used in 3D applications such as video games, proprietary 3D CAD models, VRML models used to support online commerce, and 3D elements of private medical records.

Providing secure access to interactive 3D graphics has been the subject of recent research in the computer graphics community. The state of the art in this area was significantly advanced by Koller et al. [2004]. They described a rendering technique that allows interactive, secure access to 3D models by giving the user a very low-resolution proxy version of the model for interaction purposes, and delegating high-quality renderings to a remote server to which the user does not have access. The renderings of the high-resolution models replace the low-resolution renderings when the user stops manipulating the model (on "mouse up" events), thus giving a sense of full interactivity while ensuring that the user cannot examine the 3D geometry of the high-resolution model. Other approaches to secure 3D graphics include leveraging the programmability of modern graphics chips to perform decryption in the graphics hardware [Cook et al. 2005; Shi et al. 2006]. However, these systems require platforms with specialized architectures that are not as immediately and widely deployable as the remote rendering approach of Koller et al.

Remote rendering provides an immediate level of protection because the user does not have any direct access to the full 3D geometric model. Thus a malicious attacker must resort to computer vision 3D reconstruction methods using sets of rendered images to recover a high-resolution 3D model. In order to defend against such reconstruction attacks, random perturbations can be introduced into the viewing, lighting, and shading parameters of the 3D rendering process [Koller et al. 2004]. The specific nature and magnitude of these perturbations must be chosen such that they are minimally distracting to non-malicious users, while remaining substantial enough to foil automated reconstruction efforts.

The particular perturbations used in the system of Koller et al. were chosen in a relatively ad hoc manner. Although the authors experimentally demonstrate resistance to various vision-based reconstruction attacks, the perceptual effects of the perturbations on users were not analyzed. In order to mask the sudden random changes in the viewpoint and other rendering parameters, they rely on the change blindness effect inherent in their system, due to switching between the different 3D model resolutions used for interaction and rendering. We expect that a more rigorous analysis of the perceptual effect of such perturbations would have two benefits. First, it would allow a deeper understanding of the tradeoff between perturbation defenses and the security of 3D content against reconstruction attacks, thus providing specific guidance for implementers of protected remote rendering systems. Secondly, we anticipate that future solutions for secure 3D graphics (such as those including decryption and/or a trusted graphics pipeline) will still rely on forms of perturbations and image distortions in order to combat the unpreventable reconstruction attacks that are possible utilizing sets of acquired images.

This paper thus takes the first steps towards this analysis of perceptibility of perturbation defenses for secure 3D graphics. We describe three user studies that attempt to measure the user sensitivity to and performance impact of noise added to the viewpoint control interaction of the user with a 3D model during an object examination task. The results of these studies show that there exists a noise level such that the user distraction and performance is minimally affected, whereas the same noise can significantly diminish the effectiveness of 3D reconstruction attacks.

*Department of Computer Science, University of Virginia, {jz8p, koller, humper}@virginia.edu

†Department of Psychology, University of Virginia, {jzb3e, tab2v, drp}@virginia.edu

2 User Studies

The three user studies were conducted primarily on a Dell Dimension 8250 personal computer equipped with an NVIDIA 6800XT video card. The 3D models were viewed on a 19-inch display and were rotated using a Microsoft IntelliMouse 3.0. The mouse pointer speed and acceleration were left at the Windows XP default settings.

2.1 Study 1: Preferred object rotation speed

The purpose of Study 1 was to ascertain the rotation speeds that users prefer when examining a 3D model in careful detail (see Figure 4). A conservative value based on the preferred speed obtained in this study will be used as the default rotation speed in Studies 2 and 3. We define the rotation speed as the model orientation change (in degrees) per pixel of mouse cursor movement.

2.1.1 Users

Twenty (8 male, 12 female) college students participated in Study 1. They were either compensated with snacks or paid $4.00 for their time. All gave informed consent before participating.

2.1.2 Method and procedure

Before testing, users indicated their overall experience with video games on a 7-point Likert scale (1 = no experience to 7 = a lot of experience). They also indicated the average number of hours spent per week playing video games.

Each study participant was then instructed to imagine they were a scholar who would be carefully examining all of the details of a 3D model. To do so required rotating the model.

The users' task was to adjust the object rotation speed until the speed was reached that made the model easiest to examine. Rotations were performed by holding down the left mouse button while moving the mouse. Rotation speed could be adjusted in large increments by pressing the "[" and "]" keys (much slower and faster rotation, respectively), and in finer increments by pressing the "-" and "=" keys (slightly slower and faster rotation, respectively). After finding the most comfortable rotation speed, users moved on to the next trial by pressing the right arrow key. No time constraints were specified. The study typically took five to ten minutes to complete.

Preferred speed was measured in each of the ten trials using the method of limits. The starting speeds ranged from 0.1 to 1.0 degrees/pixel in increments of 0.1. Starting speed was randomized across trials for each user.

2.1.3 Results and discussion

Across all users, the preferred rotation speed was ($M = 0.725$ deg/pixel, $SE = 0.244$). However, one user had a mean rotation speed that was more than three standard deviations from the group mean. This outlier was excluded, and the mean preferred rotation speed became ($M = 0.477$ deg/pixel, $SE = 0.039$).

3D objects can be protected from computer vision reconstruction attacks by introducing noise perturbation during object rotation (i.e. mismatches in the mapping between mouse movement and object rotation). The detection of noise has a reciprocal relationship to rotation speed (i.e. noise becomes less detectable at higher rotation speeds), and we wanted to test under conditions in which subjects were most vulnerable to the effects of noise. Thus, we adopted a conservative (slower) object rotation speed value corresponding to the 20th percentile of the measured preferred rotation speed. This means that on 80% of the trials users had a higher preferred rotation speed. This conservative value was 0.25 degrees/pixel, and was used as the *default rotation speed* in subsequent studies.

Since video game experience and time spent per week playing video games were positively correlated, $r(18) = .79, p < 0.001$, a composite variable was created by taking an average from z-scores for each of the measures. No correlation was found between the composite measure of video game experience and preferred mean speed, $p = .25$. These results suggest that there is no relationship between video game playing and preferred speed.

2.2 Study 2: Quantifying the effects of noise on user interaction in object rotation

Higher levels of perturbation provide better security, but the consequences in usability are unknown. The purpose of Study 2 was to ascertain the effect of varying levels of perturbation on user interaction. The particular perturbation that we chose to study was noise added to the mouse-controlled rotation of the 3D model; this noise perturbation thus corresponds similarly to the view direction perturbations described in [Koller et al. 2004].

The effect of the noise perturbations was evaluated in two ways: 1) by determining user sensitivity to different levels of noise, and 2) by measuring the effects of noise on the time to rotate and select an object region. A 3D model of a cow was used in this study (see Figure 5).

2.2.1 Users

Twenty-six (9 male, 17 female) subjects participated in Study 2. They were paid $4.00 for their time. All gave informed consent prior to their participation.

2.2.2 Method and procedure

Users initially filled out the same survey on video game experience used in Study 1. The study task was to "capture" different regions of the model as quickly as possible. This was done by using the mouse to rotate the model until a yellow dot located on the model aligned with the fixed crosshairs. When this was achieved, the crosshairs turned white to indicate the end of the trial. Users were told that on some of the trials (noise trials) the rotation of the model may not exactly match their mouse movements. After each trial, users used the keyboard to answer whether they thought the model movement was distorted (yes or no) and to give a confidence rating for that response on a 6-point scale (1 = guessing to 6 = certainty). The time to complete each trial was also recorded by the experimental software.

The default rotation speed determined in Study 1, 0.25 deg/pixel, was used in Study 2. There were four different trial types; no noise and three different levels of coherent noise in the mapping between mouse movement and object rotation. More specifically, we added Perlin noise [Perlin 1985] to perturb the model rotation angle and the noise was scaled by a different coefficient at each level. The highest noise level corresponded to scaled Perlin noise varying over a range between -6 degrees and +6 degrees. The medium noise corresponded to Perlin noise within the range of ± 4 degrees, and the low noise level corresponded to ± 2 degrees. Each user completed 88 trials, with the first eight trials designated as practice. Therefore, each noise level was experienced twice in practice and 20 times for the test trials. The yellow dot appeared in a unique location on every trial of a given type. The same set of dot locations was used across trial types to facilitate comparison between noise conditions. The order of trial types was randomized across trials for each user. The study typically took twenty to twenty-five minutes to complete.

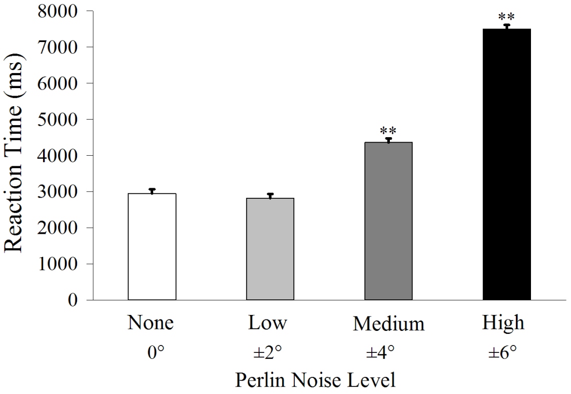


Figure 1: *Study 2. Time to complete each trial as a function of the noise level. Error bars represent one standard error of the mean. ** denotes a significant difference between conditions based on Tukey's HSD post-hoc test, $p < .001$.*

2.2.3 Results and discussion

Users' thresholds for detecting perturbations in mouse movement fell within the range of values presented. At around the medium noise level, detection was at chance (50%). Based on the psychophysical function, detection became reliable (75%) when nearing the highest noise level (see Figure 6). The data shown in Figure 6 were fit using a logistic distribution with psignifit 2.5.6 default options [Wichmann and Hill 2001].

The time to complete each trial was analyzed using a mixed-model ANOVA with trial type specified as a fixed factor and study participant as a random factor. Trials with a reaction time more than three standard deviations away from the mean for a given condition were excluded from the analysis. Approximately 4% of the total trials were excluded for this reason. Skew was corrected by using a log transform normalizing the data. There was a significant difference between noise levels, $F(3, 75.90) = 239.68, p < .001, \eta_p^2 = .91$, (see Figure 1). Medium and high levels of noise resulted in increased reaction times.

While Figure 6 indicates that participants were no better than chance at detecting the medium level of noise, Figure 1 shows that their reaction times to moderate noise were significantly impaired. The reaction time measure appears to be the more sensitive measure of noise perturbations.

A composite variable for video game experience was created using the same method as in Study 1. No correlations were found between the composite measure of video game experience and the point of subjective equality or threshold for reliable detection, $ps > .22$. In addition, no correlations were found between the composite measure of video game experience and the trial reaction times, $ps > .12$. These results indicate video game playing is not related to measures of usability in graphics security.

2.3 Study 3: Effects of noise on user reaction time

The longer reaction times found at higher noise levels in Study 2 suggest that performance is impaired at these levels. Alternately, users could have taken longer to complete particular types of trials if they were selectively spending more time assessing the possible presence of noise on those trials. Therefore, the purpose of Study 3 was to determine if asking about the presence or absence of noise had any effect on the time to complete each trial.

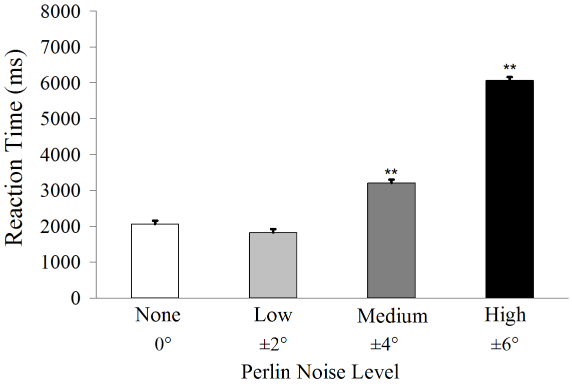


Figure 2: *Study 3. Time to complete each trial as a function of the noise level. Error bars represent one standard error of the mean. ** denotes a significant difference between conditions based on Tukey's HSD post-hoc test, $p < .001$.*

2.3.1 Users

Twenty-eight (19 male, 9 female) subjects participated in Study 3. They were paid $4.00 for their time or they received course credit. All gave informed consent prior to their participation.

2.3.2 Method and procedure

The method and procedure of Study 3 were identical to Study 2, except users did not give a yes/no response or confidence rating about noise. They were simply instructed to complete each trial as quickly as possible. The study typically took fifteen to twenty minutes to complete. The default rotation speed of 0.25 degrees/pixel was used again in this study.

2.3.3 Results and discussion

The same statistical analysis as in Study 2 was conducted on these data. About 4% of the trials were excluded because they were classified as outliers. A log transform was conducted to normalize the data. As in Study 2, there was a significant difference between noise levels, $F(3, 75.57) = 252.73, p < .001, \eta_p^2 = .91$, which was again limited to the medium and high noise levels.

The reaction time results from Study 3 mirror those from Study 2 (see Figure 2), although the reaction times at all noise levels are about 1 second lower than in the prior study. This uniform decrease in reaction time suggests that asking about assessment of noise in Study 2 did not have selective effects on reaction times.

A composite variable for video game experience was created using the same method as in Studies 1 and 2. As in Study 2, no correlations were found between the composite measure of video game experience and the reaction times, $ps > .60$. This provides converging evidence that video game playing is not related to measures of usability in graphics security.

3 Reconstruction attacks

The four interaction noise levels introduced in Study 2 correspond to varying magnitudes of noise perturbation defenses against malicious 3D reconstruction attacks. To evaluate the effectiveness of such defenses, we can apply 3D reconstruction algorithms to sets of images acquired at the various noise levels, and measure the quality of the resulting reconstructed models.

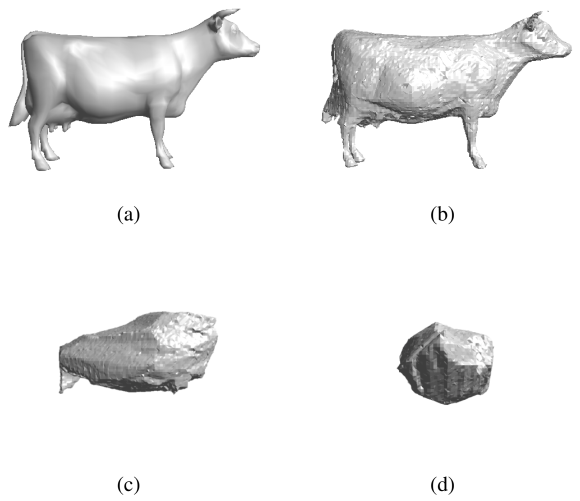


Figure 3: *Shape-from-silhouette 3D reconstruction attack results. (a) original 3D model ($E = 0.0$), (b) No noise added ($E = 0.0047$), (c) Low noise level ($E = 0.022$), (d) Medium noise level ($E = 0.032$). Error values (E) measure the mean surface distance from the original model.*

Example results of such a simulated reconstruction attack are shown in Figure 3. In this case we have executed a shape-from-silhouette reconstruction algorithm on sets of 50 images harvested from a variety of viewpoints around the 3D model, for each of three different interaction noise levels (no noise, low noise, and medium noise). Shape-from-silhouette is a well-studied technique for extracting a 3D model from a set of images by segmenting the object pixels from the image backgrounds, and then computing the shape of the intersection of their extended silhouettes [Slabaugh et al. 2001].

Renderings of the resulting 3D models reconstructed with the shape-from-silhouette approach as well as the mean surface deviations from the original model (as a fraction of the bounding box dimensions, as computed with the Metro tool [Cignoni et al. 1998]) are given in Figure 3. With no noise perturbations affecting the images, a relatively accurate reconstruction of the original model is possible (Figure 3(b)). However, at the low noise level, the shape of the reconstruction output is barely recognizable (Figure 3(c)), and at the medium noise level the output is amorphous (figure 3(d)). Niem [1997] performed an error analysis of silhouette-based modeling techniques and showed the linear relationship between error in the estimation of the view position and error in the resulting reconstruction.

4 Conclusion

In this paper, we have taken the first steps towards analyzing the perturbation defenses that are a fundamental part of state of the art protected 3D graphics systems. Our series of user studies measured the perceptibility and performance impact of noise added to the mouse control of viewpoint position at several different noise levels. The results of Study 2 and Study 3 show that users are sensitive to both the medium and high noise levels, with impaired performance on the user task as measured by reaction time. At the low noise level, however, users do not experience a significant difference in performance, and cannot easily detect the noise.

As evidenced by Figure 3(c), however, the low noise level of perturbation defense is able to significantly degrade the quality of results of a simulated reconstruction attack. Thus, the low noise level used in this paper represents a “sweet spot” in the tradeoff between magnitude of perturbation defenses for 3D graphics security, and minimizing the perceptual impact of the defenses on user performance. Though the direct applicability of our immediate results to other usage scenarios and protected graphics systems is limited, they do provide evidence and optimism that an appropriate balance can be found between security and usability for interactive 3D graphics content.

In our future work, we will continue to investigate the applicability of perturbation defenses in protected rendering systems for 3D graphics. We will extend our empirical study to include more sophisticated reconstruction attacks beyond the simple shape-from-silhouette attack used herein, by evaluating hybrid computer vision approaches that include photometric stereo and other contemporary techniques for effective 3D reconstruction. Additionally, we will expand the scope of perturbation defenses that we examine to include distortions such as image-space warping, and variations in the lighting and shading parameters of the rendering.

Acknowledgements

This work was supported in part by DARPA (grant NBCH-1050023 to Dennis Proffitt) and the National Science Foundation (grant IIS-0535118). We thank David Luebke and Jeanine Stefanucci for their intellectual contributions, and Allison Brennan and Amelia Swafford for their assistance with data collection.

References

CIGNONI, P., ROCCHINI, C., AND SCOPIGNO, R. 1998. Metro: measuring error on simplified surfaces. *Computer Graphics Forum 17*, 2, 167–174.

COOK, D. L., BARATTO, R., AND KEROMYTIS, A. D. 2005. Remotely keyed cryptographics - secure remote display access using (mostly) untrusted hardware. In *Proceedings of ICICS*, LNCS 3783, 363–375.

KOLLER, D., AND LEVOY, M. 2005. Protecting 3d graphics content. *Communications of the ACM 48*, 6, 74–80.

KOLLER, D., TURITZIN, M., LEVOY, M., TARINI, M., CROCCIA, G., CIGNONI, P., AND SCOPIGNO, R. 2004. Protected interactive 3d graphics via remote rendering. *ACM Transactions on Graphics 23*, 3, 695–703.

NIEM, W. 1997. Error analysis for silhouette-based 3d shape estimation from multiple views. In *International Workshop on Synthetic-Natural Hybrid Coding and 3D Imaging*.

PERLIN, K. 1985. An image synthesizer. *Computer Graphics 19*, 3 (July), 287–296.

SHI, W., LEE, H.-H. S., YOO, R. M., AND BOLDYREVA, A. 2006. A digital rights enabled graphics processing system. In *Proceedings of Graphics Hardware*, 17–26.

SLABAUGH, G., CULBERTSON, B., MALZBENDER, T., AND SCHAFER, R. 2001. A survey of methods for volumetric scene reconstruction from photographs. In *Proc. of the Joint IEEE TCVG and Eurographics Workshop (VolumeGraphics-01)*, Springer-Verlag, 81–100.

WICHMANN, F., AND HILL, N. 2001. The psychometric function: I. fitting, sampling and goodness-of-fit. *Perception and Psychophysics 63*, 8, 1293–1313.

Quantifying Usability in Secure Graphics: Assessing the User Costs of Protecting 3D Content

Zhu, Bakdash, Koller, Banton, Proffitt, and Humphreys

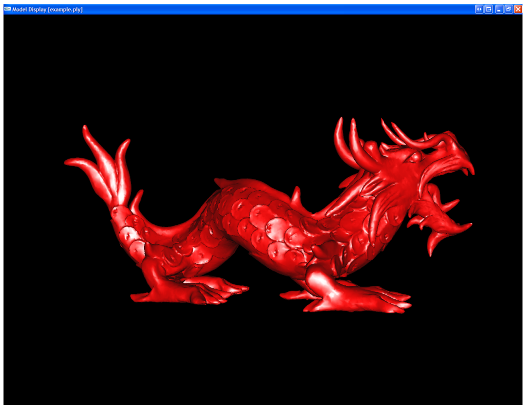
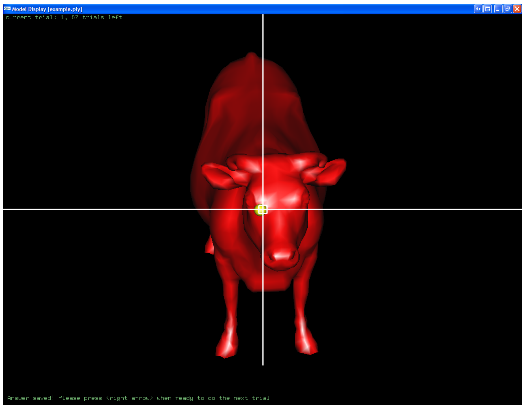


Figure 4: *The 3D model used in Study 1.*



current trial: 1, 87 trials left

Answer saved! Please press <right arrow> when ready to do the next trial

Figure 5: *The 3D model used in Studies 2 and 3. The task was to align the yellow dot with the crosshairs.*

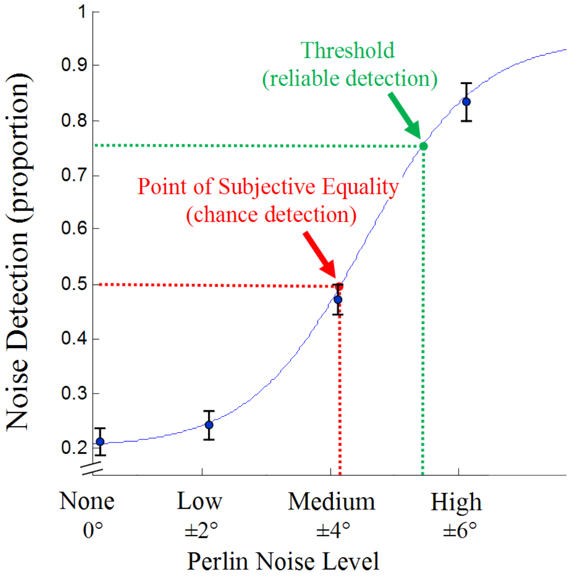


Figure 6: *Study 2. The psychophysical function for detecting noise. The y-axis represents the proportion of yes and no responses weighted by confidence ratings. The x-axis represents the four different levels of noise. Chance detection (point of subjective equality) is shown in red and reliable detection in green. Error bars represent one standard error of the mean.*