

Available from: http://dx.doi.org/10.1109/CMSP.2011.172

Copyright © 2011 IEEE.

This is the author's version of the work. It is posted here with the permission of the publisher for your personal use. No further distribution is permitted. If your library has a subscription to these conference proceedings, you may also be able to access the published version via the library catalogue.
How real is reality?
A perceptually motivated system for quantifying visual realism in digital images

Norman Wang
Faculty of Life and Social Sciences
Swinburne University of Technology
Melbourne, Australia
571060x@swin.edu.au

Wendy Doube
Faculty of Life and Social Sciences
Swinburne University of Technology
Melbourne, Australia
wdoube@swin.edu.au

Abstract—Contemporary video games are generally designed with the aim of engendering a state of immersion in their players. A sense of ‘presence’ is considered to be a prerequisite of total immersion, the ultimate level of immersion, and can be directly related to the degree of realism in the simulation presented in a videogame. However, continuously increasing the degree of visual realism in an image could cause adverse effects such as fault amplification and more importantly, emotions like fear and dread of the uncanny. It can also substantially raise the development costs of a game. Consequently, videogame developers would benefit greatly from knowing the threshold beyond which increased realism is counter-productive.

This paper details the development of an image processing system that measures the degree of realism of an image using three methods - gradient variance, color variance and shadow softness - derived from existing theories and practices in perceptual psychology, photoforensics and image processing. It then discusses testing of the system on a variety of photographic and video game images.

Keywords—video game; photorealism; immersion; presence; image classification; image processing; shadow segmentation; virtual reality; realism; uncanny valley; game studies

I. INTRODUCTION

Videogame production is rapidly becoming one of the largest of the global digital entertainment industries, with titles such as Grand Theft Auto IV, Halo 3 and Call of Duty 4: Modern Warfare amongst the most economically successful entertainment products of all time [1-3]. The growing prominence of videogames as an industry has motivated game scholars to investigate theories and devise means for maximizing player gratification [4].

The theory of immersion is frequently referred to in discussions on engineering positive gameplay experience [5, 6]. The deepest level of immersion is total immersion, in which players feel as if they are physically ‘present’ in the virtual world [5, 7, 8]. Sense of presence is proportional to the degree of visual realism in the virtual environment [7, 9, 10]. Consequently immersion in videogames can be enhanced by higher degrees of visual realism.

However, an increase in realism can cause adverse effects [11]. Firstly, according to the theory of the uncanny valley [12, 13], continual increases in realism could incite potentially negative emotional responses like fear and dread and amplify otherwise unnoticeable faults in the virtual environment when objects in the virtual world become almost indistinguishable from their counterparts in reality [9, 11]. Secondly, from an economic perspective, higher levels of visual realism greatly inflate the cost of game development, often to a level that is almost prohibitive to all but the largest developers [14, 15].

This paper describes the construction of a software system which can empirically measure visual realism in images using techniques gleaned from the fields of perceptual psychology [11, 16], photoforensics [17, 18] and computer vision [19]. The aims of the system are:

• To compare the levels of visual realism of two games, or two versions of the same game with differing visual qualities, in order to engineer realism without resorting to building higher quality art assets.
• To support further research into player responses to two games with identical gameplay but with differing levels of visual realism in order to better understand the relationship between realism, immersion and player response.

II. MEASURING VISUAL REALISM

Humans appear to share a universal perception of ‘visual reality’ [16, 20]. This universal response could be innate [21] or it could be caused by frequent exposure to images with similar qualities in the real world, i.e. surfaces in reality are very rarely ‘perfect’ – they usually display some irregularities, or could be covered in grime and dirt [16]. The latter contention has been supported in studies that showed that visual perception can be malleable and prone to external influences during the development of the brain and associated cognitive functions [22]. Regardless of the cause, it can be argued that this apparent perceptual universalism is the result of a consistent and external reality and thus it should be possible to measure visual realism.

Most existing systems capable of quantifying characteristics of realism in images are in the field of photoforensics and image processing. Many of these techniques for image classification were unsuitable for use in this project because they are a) too obscure and cannot be generalized to a large range of images [23-26] or b) based entirely on statistical modeling of characteristics beyond human perception [17, 19, 27-33].

In contrast, the field of perceptual psychology contains ample literature pertaining to human vision and measurement
of image characteristics that are perceivable, in particular, characteristics that are perceived as ‘real’ [11]. These measurements are generally derived from primarily qualitative research methods that are not readily transferable to the construction of a software tool. However, some of them are also represented in technical systems developed in photoforensics and image processing and therefore can be quantified.

Three quantifiable and perceivable characteristics of realism were identified in the fields of photoforensics [27], perceptual psychology [16] and image processing [19]: image roughness; shadow softness; and color variance. The techniques to measure them were developed with reference to their relevance to human visual cognitive functions.

III. MEASURING CHARACTERISTICS OF REALISM

The measurements for each of the three characteristics were designed to produce a standard numeric value between 0 and 1.

A. Surface Roughness Measure

Images appear more realistic when the surfaces of their objects are perceived to be rough or uneven. Conversely, they appear less realistic when the surfaces of their objects appear smooth. In Figure 1, the right hand image has a rougher surface and is considered more realistic [16].

The surface roughness measure developed in this study is based on the Laplacian of the input image, which represents the degree of variance in the intensity of a pixel in relation to all adjacent pixels, also referred to as gradient image [17].

Let \( I \) denote an image, its Laplacian image \( L(\text{Laplace}_I) \) is:

\[
\text{Laplace}_I = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}
\]

The Laplacian image is extracted with an aperture size of 3 which is, in effect, the sum of the second-order 3 x 3 Sobel derivatives in the x and y directions with no scaling operations.

The roughness measure of the image, \( M_{\text{Roughness}} \), is the Gini coefficient of the 30 bin histogram of the Laplacian image, \( H_{\text{Laplace}} \). The Gini coefficient measures the uniformity in statistical distributions and is commonly used to evaluate the degree of inequality within a set of data [34, 35]. Let \( L \) denote the Lorenz curve of \( H_{\text{Laplace}}. M_{\text{Roughness}} \) is obtained as follows:

\[
M_{\text{Roughness}} = 1 - 2 \int_0^1 L(H_{\text{Laplace}}) \, dH
\]

B. Color Variance Measure

The color variance measure developed in this study is based on the methods used by Cutzu et al. [19] and investigates the possibility that game images could differ from typical photographic image in the number of unique colors. The color variance measure is the variance in color within the image calculated by normalizing the number of unique colors against the number of pixels in the picture.

Let \( I \) be an image over RGB space and \( N \) be the number of pixels in \( I \) and \( I_C \) denote set of colors where each element of the set is a unique tuple of RGB values:

\[
I_C = \{(R_1, G_1, B_1), (R_2, G_2, B_2), ..., (R_n, G_n, B_n)\}
\]

The color variance measure is a scalar value defined as:

\[
M_{\text{Colour}} = \frac{|I_C|}{N}
\]

C. Shadow Softness Measure

Images in which objects project hard shadows under the illumination of strong, directional light are perceived as less ‘real’ than images in which soft shadows are projected under normal diffused illumination, irrespective of whether the images are photographic or whether they are digitally rendered [16]. In Figure 2, the images display shadows that are graded from extremely hard in the top left to extremely soft in the bottom right. The images with hard shadows are perceived as the most unrealistic while images with an intermediate level of shadow softness (5.21) are perceived as the most realistic [16].

The shadow softness measure is obtained by calculating the histogram of the image and normalizing the number of pixels within each bin to the number of pixels in the image.
The shadow softness measure developed for this study measures the amount of ‘gradation’ at the edge of shadow regions. The shadow region segmentation technique is based on the observation that shadow pixels in an HSV image typically have low value, high hue and high saturation [36, 37] and makes use of the Otsu thresholding function [38] over the HSV color channels of the input image as developed by Su et al. [36] and Tsai [39].

Let $H, S, V$ respectively denote the hue, saturation and value channel of the image over HSV color space, the spectral ratio image $R$ is defined as:

$$R_{(x,y)} = \frac{H_{(x,y)}}{V_{(x,y)} + 1}$$

The shadow region is a set of $(x,y)$ coordinates selected using the following qualifications:

$$\{(R_{(x,y)} > T_R) \land (B_{(x,y)} < T_B) \land (S_{(x,y)} > T_S)\}$$

Where $T_R, T_B, T_S$ respectively denote the Otsu threshold of the spectral ratio image $R$, the input image over RGB color space $B$ of and the saturation channel $S$. The final measure of shadow softness is defined as follows:

$$M_{\text{Shadow}} = \frac{\left|(R_{(x,y)} > T_R \land B_{(x,y)} < T_B \cdot (1 - \Delta)) \land S_{(x,y)} > T_S\right|}{\left|(R_{(x,y)} > T_R \land B_{(x,y)} < T_B \cdot (1 + \Delta)) \land S_{(x,y)} > T_S\right|}$$

Where $\Delta$ stands for the arbitrary displacement from the ideal Otsu threshold value. In this case, a $\pm 10\%$ displacement was applied to the Otsu threshold to account for the gradation of shadow from the ideal shadow edge.

IV. CALCULATING REALISM METRICS

In order to formulate realism metrics for the level of visual realism in an image, the three image measures for that image (discussed in Section III) are compared against three reference measurement sets, $R_{\text{Roughness}}, R_{\text{Shadow}}$ and $R_{\text{Color}}$ based on a library of 99,837 reference photographs collected by Allan and Verbeek [40].

A. Reference Measurements

Each of the reference sets contains the three characteristic measurements from $n$ numbers of photographs in the reference library. They are defined as follows:

$$R_{\text{Roughness}} = \{M_{1}^{\text{Roughness}}, M_{2}^{\text{Roughness}}, ..., M_{n}^{\text{Roughness}}\}$$

$$R_{\text{Color}} = \{M_{1}^{\text{Color}}, M_{2}^{\text{Color}}, ..., M_{n}^{\text{Color}}\}$$

$$R_{\text{Shadow}} = \{M_{1}^{\text{Shadow}}, M_{2}^{\text{Shadow}}, ..., M_{n}^{\text{Shadow}}\}$$

B. The Realism Metrics

The realism metrics, $\text{Metric}_{\text{Roughness}}, \text{Metric}_{\text{Colour}}$ and $\text{Metric}_{\text{Shadow}}$ of any given image is the relative standing of $M_{\text{Roughness}}, M_{\text{Colour}}$ and $M_{\text{Shadow}}$ of that image to the reference measurement set $R_{\text{Roughness}}, R_{\text{Colour}}$ and $R_{\text{Shadow}}$. It is defined as follows:

$$\text{Metric}_{\text{Roughness}} = \frac{|M_{\text{Roughness}} - \mu(R_{\text{Roughness}})|}{\sigma(R_{\text{Roughness}})}$$

$$\text{Metric}_{\text{Colour}} = \frac{|M_{\text{Colour}} - \mu(R_{\text{Colour}})|}{\sigma(R_{\text{Colour}})}$$

$$\text{Metric}_{\text{Shadow}} = \frac{|M_{\text{Shadow}} - \mu(R_{\text{Shadow}})|}{\sigma(R_{\text{Shadow}})}$$

The composite realism metric is defined as the sum of the three separate metrics.

$$\text{Metric}_{\text{Composite}} = \text{Metric}_{\text{Roughness}} + \text{Metric}_{\text{Colour}} + \text{Metric}_{\text{Shadow}}$$

In all cases, a more realistic image is expected to have smaller metrics.

V. EXPERIMENT

A software system was constructed to implement the measures discussed in Section IV. The system can process multiple images in a single execution. The aims of the experiment were to verify that the software system could a) calculate reasonable measures for three characteristics and a composite metric for realism in video game images; and b) use those measures and metric to differentiate the levels of realism of two similar video game images.

A. Experiment Method

The reference set of photographs was processed by the system to establish benchmarks for the three characteristic metrics and the composite metric. A test data set of 168 game images with a variety of visual styles and levels of realism was processed to produce equivalent measures.

Firstly, the distribution of the data from the games was compared against that of the reference measurement set to identify any fundamental differences between game images and photographic images and to confirm that the system’s calculations were reasonable. Images with higher levels of visual realism were expected to score a measurement closer to the reference measurement and therefore have a lower $\text{Metric}_{\text{Composite}}$.

Secondly, two sets of six game images were selected for an in-depth comparison. The first set was selected quantitatively based on the composite metrics of the image and the second set was selected qualitatively on a) subjective evaluation of realism; and b) date of production. Given that increases in technology allows for increased levels of realism in computer generated images [28, 29], it is expected that visual realism in an image is proportional to its production
year. The quantitatively selected images were then qualitatively evaluated and vice versa to determine whether the metric reflected subjective perception of realism.

B. Test Data

The test data was a set of 168 game images sourced from 42 games released between 1993 and 2010, including one game that, at the time of writing this paper, is yet to be released. These images were selected based on three criteria:

1. Minimal special effects such as explosions or muzzle flash, as these elements are very bright and lacking in texture and would skew the roughness and color measurement of the image.
2. No HUD (Heads Up Display) or other overlays as they create sharp edge transitions and contain color that is not normally present in the scene, thereby affecting the roughness and color variance measurement.
3. A large range of game genres with a variety of visual styles to ensure diversity in the test data.

Ideally, the test game image dataset would have a subset of images depicting a variety of equivalent scenes for each of the game titles used as sources of game images. However, difficulties involved in sourcing a large number of applicable images from 42 games restricted the number of game images to roughly 4 images per game in order to prevent games with a wide availability of images from skewing the test dataset.

VI. RESULTS

A. Image Measurement Output

The intra-set variance of the measurements for both the reference and test data sets is displayed in Table I and represented in boxplot form (Figure 3):

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Max</th>
<th>Median</th>
<th>Min</th>
<th>( \mu )</th>
<th>( \sigma )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( R_{\text{Roughness}} )</td>
<td>0.956</td>
<td>0.790</td>
<td>0.575</td>
<td>0.787</td>
<td>0.068</td>
</tr>
<tr>
<td>( M_{\text{Roughness}} )</td>
<td>0.950</td>
<td>0.852</td>
<td>0.672</td>
<td>0.845</td>
<td>0.056</td>
</tr>
<tr>
<td>( R_{\text{Color}} )</td>
<td>0.667</td>
<td>0.294</td>
<td>0.050</td>
<td>0.306</td>
<td>0.132</td>
</tr>
<tr>
<td>( M_{\text{Color}} )</td>
<td>0.759</td>
<td>0.311</td>
<td>0.066</td>
<td>0.333</td>
<td>0.152</td>
</tr>
<tr>
<td>( R_{\text{Shadow}} )</td>
<td>1.208</td>
<td>0.900</td>
<td>0.045</td>
<td>0.896</td>
<td>0.202</td>
</tr>
<tr>
<td>( M_{\text{Shadow}} )</td>
<td>1.133</td>
<td>0.932</td>
<td>0.676</td>
<td>0.931</td>
<td>0.122</td>
</tr>
</tbody>
</table>

The boxplots illustrate the similarities between the distribution of measurements for photographic and game images. However, they also indicate some fundamental differences.

1) Surface Roughness Measure

The roughness measure set shows that 86.8% of the game images produced measures above the mean value of the images in the reference set. As a smaller \( M_{\text{Roughness}} \) denotes a more evenly distributed pixel variance and lower levels of realism, this indicates that the system recognized the surfaces in the game images as having less varied degrees of roughness than photographs. Furthermore, blurring effects such as motion blur and depth of field were associated with higher roughness metrics even when they appeared highly realistic subjectively. Elements such as fog, clouds, sky and ocean had a similar effect. The system output for the surface roughness measurement for one image is shown in Figure 4.

2) Color Variance Measure

The distribution of color measures for game images was extremely similar to that of the reference set but slightly wider. This implies that the measure was not an effective technique for differentiating game images from photographs.

3) Shadow Softness Measure

The shadow softness measure for 65.3% of the game images was above that of the reference set mean indicating slightly lower levels of realism. However, the reference set measures contained a large number of statistical outliers as well as an unexpected outcome where approximately 15% of the images produced measures greater than 1. This is discussed in further detail in Section VII.C. After removing the outliers, the mean of the reference set was reduced from 0.8969 to 0.8771 and the percentage of game images above the reference mean increased from 65.3% to 77.5% indicating that the system recognized game images as having harder shadows than those of photographic images. The system output for the shadow roughness measurement for one image is shown in Figure 5.
B. Realism Metrics

As described in Section IV, this metric represents the deviation of the three image measurements from the mean of their respective reference set with a lower metric value denoting a more realistic image. The descriptive statistics for the realism metrics obtained from the test images are displayed in Table II and represented in boxplot form in Figure 6:

<table>
<thead>
<tr>
<th>Metric</th>
<th>Max</th>
<th>Median</th>
<th>Min</th>
<th>μ</th>
<th>σ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Roughness</td>
<td>2.439</td>
<td>1.078</td>
<td>0.004</td>
<td>1.066</td>
<td>0.573</td>
</tr>
<tr>
<td>Color</td>
<td>2.585</td>
<td>0.788</td>
<td>0.008</td>
<td>0.926</td>
<td>0.714</td>
</tr>
<tr>
<td>Shadow</td>
<td>1.772</td>
<td>0.582</td>
<td>0.001</td>
<td>0.697</td>
<td>0.744</td>
</tr>
</tbody>
</table>

To investigate the effectiveness of the metric in differentiating images of varying levels of realism, the images were sorted in descending order of realism according to Metric\textsubscript{Composite} of the image. This sorted list was divided into six evenly spaced intervals and the first image in each interval was selected for comparison. The ordering of the six images was analyzed by qualitative assessments of realism.

The metric values for the six images are illustrated in the boxplots in Figure 7.

Next, six images were qualitatively selected based on subjective assessment of their apparent level of visual realism and the release date of the game. The ordering of the images was then compared with their realism metrics, which are displayed on the metric boxplot in Figure 10. The boxplot shows a correlation between the order rankings for roughness and shadow metrics and qualitative evaluation of the realism of the images.

To provide a visual representation of the realism metrics, Figure 6 illustrates the distribution of game image realism metrics.
VII. DISCUSSION

A. Reference Measurements and Game Image Measurements

The overall distribution of the measurements (Figure 3) shows that the system quantified game images as having similar but lower levels of realism in comparison to photographs for two image characteristics:

a) Less variance in surface roughness
b) Harder or softer shadows

These findings are consistent with those of Pan et al. [17] and Rademacher et al [16] and confirm some fundamental differences between game and photographic images. They also indicate that the system functioned appropriately overall in measuring two characteristics of realism in images.

However there was no noticeable overall difference between game images and photographs in their richness of color. This could be because color variance in images is predicated more on the scene they depict than on the means of their creation. For example, images depicting urban landscapes and indoor scenes tend to have less color richness than images of natural scenes like forest landscapes, regardless of whether they are from a game or from a photograph.

B. Image Ranking by Composite Metric

The composite metric ordering was expected to be largely predicated on year of production, working on the assumption that realism has increased with time. However, two of the six images selected by composite metric (Figure 8) could be considered ‘stylized’, or designed to exemplify an art style that deliberately deviates from realism by conveying a more abstract or personal view [41]. This may be achieved with techniques such as minimal use of photographic textures and exaggerated color.

The composite metric ranked the test images by year of production except in the cases of the two stylized images which were classified as less real than other more realistic images produced in earlier years.

The roughness measure ranked all images in the order of the composite metric, indicating that gradient variance provides a strong indicator of realism. The color variance metric followed the same order with all images except for image 2 which depicts an indoor scene with a limited number of objects – a zombie, some rubble and a window. This scene could be expected to have a dearth of color. The shadow metric separated the images into two distinct groups of rankings.

Shadow softness does not necessarily vary with stylization and therefore was not expected to be a factor in ordering the stylized images. The images with metrics indicating less realistic shadows are either silhouetted against a sky or an unrealistic indoor scene.

In summary, the image rankings by composite metric, together with their component metrics, largely reflected qualitative evaluations of realism of the images.

C. Image Ranking by Qualitative Selection

In general, the roughness and shadow metrics ranked the qualitatively selected images from high to low by production year, indicating a higher level of visual realism in more recent games. These metrics consistently rated Crysis 2 as the most visually realistic in the test set. This result was expected, as a) Crysis 2 is the most recent game and b) the images sourced from Crysis 2 were subjectively as well as empirically evaluated to be the most visually realistic in the test set (image 1 in Figure 8 and Figure 9).

The roughness metric accurately ranked all except one image according to release date and apparent level of visual realism. In the case of the anomaly (Figure 11), the large number of sharp transitions produced more evenly distributed Laplacian values, resulting in a measure typically expected from a more visually sophisticated image. A method that could correct this behavior is proposed in Section VIII.B.

In contrast, the shadow metric was able to accurately place this image in the rankings according to its release date and apparent level of visual realism. The shadow metric accurately placed all but one qualitatively selected image (Figure 10) in accordance to its production date. That image, sourced from Silent Hill 3, illustrates a higher level of detail
and lighting technique than usually found in games of its year of production.

The color variance metric did not reflect the qualitative selection criteria and was largely determined by type of scene with outdoor versus indoor, rural versus urban being factors that confounded production date.

Overall, the metrics were highly effective. The roughness and shadow metrics ordered game images according to qualitative assessment of visual realism. The color metric was highly dependent on the type of scene and the objects contained within the scene. Although contributing to the overall composite metric, it did not accurately distinguish levels of realism.

The following section discusses anomalies which prevented the system from accurately measuring a small number of images.

D. Shadow Softness Reference Set Anomalies

As described in Section III.C, $M_{Shadow}$ is the ratio between the cardinality of two sets of pixels at two different thresholds and the core of the measure is the ratio between the number of pixels in the set $\{B_{(x,y)} < T_B \cdot (1 - \Delta)\}$ and the set $\{B_{(x,y)} < T_B \cdot (1 + \Delta)\}$. As the latter set is always inclusive of the former, the calculation should always yield a result between 0 and 1. A result larger than 1 indicates that the image contains more shadow pixels in the lower threshold than the higher threshold.

This phenomenon occurred only in images with very limited color variation, typically depicting a small object against a marginally varied background, e.g. images of planes or birds against the sky.

These images tend to have a small number of unique colors (<10,000) and consequently a small $M_{Colour}$ (< 0.1).

As the shadow measure was implemented through the OpenCV image processing library, this phenomenon could have two causes:

1. An error in OpenCV’s implementation of thresholding functions, where the $\pm 10\%$ deviation from $T_B$ does not provide adequate distinction between the two sets of shadow pixels.
2. These images do not contain sufficient variance in the hue and saturation color channel to allow for a meaningful computation.

In either case, this problem could be resolved through the use a more robust reference image filter, which is detailed in Section VIII.A.

VIII. Future Works

Although the system performed largely as intended, the anomalies discussed in Section VII.B and Section VII.C suggest that there is scope for further enhancement.

A. Expanded Reference Image Filter

As discussed in Section VII.C, images with insufficient information in the hue and saturation channel created errors in the shadow softness measurement. Currently the pre-processor filters out images based on the number of unique colors in that image, where a count of $\leq 256$ denotes a grey-scale image. A more robust filter should classify input images in accordance with the variance in the hue and saturation channel of the image and should help minimize the number of outliers in the reference shadow measurements.

B. Bhattacharyya Coefficient Roughness Metric

The roughness measure of an image is the Gini coefficient of the Laplacian histogram and does not account for where the unevenness in distribution occurs for an image. As outlined in Section VII.B, the ideal roughness metric should be able to differentiate whether the image is excessively smooth or excessively rough. This could be achieved through the use of Bhattacharyya coefficient, which measures the degree of similarity between two distributions, typically seen in comparative histograms analysis.

C. Additional Measures

The software system has an architecture that enables the addition of other measures of realism, such as:

a) Pattern detection for differentiating images with large regions of repetition, which are characteristically unrealistic

b) Color saturation measures for the ratio of highly saturated pixels in an image where a higher ratio could indicate a lower level of realism

D. Studies in Immersion

The software system is now functioning at a level in which it can be used in experiments that evaluate game player responses to realism in game images. A system has been designed to enable it to be used in dynamic video as
opposed to still images and this will be documented in another paper.

IX. CONCLUSION

This paper has demonstrated how visual realism, a somewhat abstract concept, can be measured and quantified. It describes the rationale and testing of a software system which can empirically measure three realism benchmarks of any digital image: the surface roughness of objects in the scene; the variance of color in the scene; and the softness of the shadow in the scene. Two of the three realism metrics (roughness and shadow) were able to accurately differentiate images with varying degrees of visual realism.

The next phase in this investigation of immersion and realism is to employ the system in experiments that measure player response at varying levels of visual realism. In particular, the investigation will compare levels of immersion with levels of realism with the aim of identifying the threshold of the uncanny valley.

The ability to quantify the degree of realism in specific visual characteristics and relate them to player sense of immersion should enable developers to better focus their resources. This could be an important step in game studies that could direct existing development practices to more financially viable models. Independent developers would be able to engineer realism in games without needing a substantial development budget. The system described in this paper could also have relevance beyond the videogame industry.

ACKNOWLEDGMENTS

The authors would like to acknowledge the following individuals for their assistance and support throughout the creation of this paper:

Mr. James Bonner – For editorial input
Dr. Clinton Woodward – For editorial input and technical assistance
Mr. Rajash Vasa – For assistance with statistical measurements and metrics

REFERENCES


