Technical Section

Smooth shadow boundaries with exponentially warped Gaussian filtering

Jesús Gumbau\textsuperscript{a,*}, Mateu Sbert\textsuperscript{b}, László Szirmay-Kalos\textsuperscript{c}, Miguel Chover\textsuperscript{a}, Carlos González\textsuperscript{a}

\textsuperscript{a}Universitat Jaume I, Castellón, Spain
\textsuperscript{b}Universitat de Girona, Girona, Spain
\textsuperscript{c}BME, Budapest, Hungary

\textbf{A B S T R A C T}

Shadow mapping is widely used in computer graphics for efficiently rendering shadows in real-time applications. Shadow maps cannot be filtered as regular textures, thus their limited resolution can cause severe shadow map discretization artifacts in the rendered images. To solve this problem, several techniques have been proposed, including variance shadow maps (VSM) and exponential shadow maps (ESM). However, these techniques introduce different kinds of “light leaking” artifacts, which are clearly visible in moderately complex scenes. In this paper we propose a new statistical filtering method that approximates the cumulative distribution function (CDF) of depth values by a Gaussian CDF instead of bounding it with Chebyshev Inequality. This approximation significantly reduces “light leaks” and has similar performance and storage requirements compared to the original variance shadow map method. We also show that the combination of this technique with an exponential warp allows us to further reduce the remaining shadowing artifacts from the rendered image.

\section{Introduction}

Shadow mapping is a popular and efficient technique for solving the shadowing problem in interactive applications. A shadow map approximately represents the light occluding geometry as depth or distance values sampled on a two-dimensional grid. This sampled depth function is then used to reconstruct the distance between the light source and a point to be shaded in order to decide whether or not the point is behind the occluding geometry, i.e. it is in shadow.

The shadow test can also be imagined as a step like \textit{visibility function} \( v(z_r) = \mathbb{1}(z_0 - z_r) \) that is 0 if the occluder distance \( z_0 \) from the light source to the occluder surface is smaller than receiver distance \( z_r \) measured between the light source and the point to be shaded, and 1 otherwise. However, the light occluder distance is known only in the centers of the shadow map texels, which leads to shadow map sampling artifacts during the visualization process. In order to avoid this issue, the occluding surface (or the visibility function) must be reconstructed and turned into a continuous function. However, as the geometry of the occluder may involve high frequency variations, the occluder distance function can only be approximately reconstructed, and high frequency components may distort the reconstructed signal even at low frequencies, which leads to the well known phenomenon of shadow aliasing.

Linear signal theory has a solution for the aliasing problem, which is based on the usage of low-pass filters in order to eliminate high frequencies that are above the Nyquist limit. However, shadow mapping is a non-linear operation as it contains a comparison operation represented by the \textit{step function}. As a consequence, filtering the depth values directly before performing the comparison operation would result in averaged filtered depth values, which would lead to incorrectly calculated shadow boundaries. The problem of the non-linear step function is solved by simply delaying the filtering operation after performing the comparison, which is the basic idea of \textit{percentage closer filtering} (PCF). This approach, which is able to present artifact-free correct anti-aliased shadows, has two major drawbacks. First, the filtering can only be performed when the distance of the shaded point is available, which requires it to be executed separately for every shaded point. Second, many shadow map texture fetch operations must be performed in order to achieve smooth anti-aliased shadow boundaries.

The final goal for solving this problem would be the ability of merging the filtering operations that are executed separately for
each shaded point by PCF into a single texture filtering step. We wish to perform filtering as a pre-process before the rendering stage. This can be solved by storing additional information into the shadow maps so that the distribution of depths can be effectively approximated. The information stored in the shadow map must be GPU-friendly, which means that the visibility function can be evaluated from the filtered data with texture fetching modes supported by the graphics hardware, which are restricted to the linear interpolation operation.

Existing techniques aiming at pre-filtering the shadow map can be classified into two main categories, those that transform the problem into a domain where linear signal processing becomes feasible, and those that use non-linear filtering operations based on statistical analysis. Both families of methods allow for efficiently filtering the shadow map on the graphics hardware, but they also introduce artifacts commonly known as “light leaking”, which incorrectly alter the shadow intensity.

The method presented in this paper, which is an extended version of [1], belongs to the category of statistical filtering. It is based on approximating the probability that the shaded point passes the depth test. This probability is obtained from the approximation of the cumulative distribution function of depths with a Gaussian CDF instead of bounding it with the Chebyshev Inequality (as happens in VSM) or instead of approximating the step function with an exponential function (as in ESM). The two moments of the depth’s distribution are used to construct a Gaussian CDF. This approach is capable of highly reducing the “light leaking” artifacts present in existing techniques, or even eliminating it for moderately complex scenes, with neither penalty of performance nor increased storage cost. In addition to [1], we present an extension to our technique that builds upon the pure Gaussian approach and allows us to further minimize the appearance “light leaking” artifacts from the rendered image.

There are several works in the literature for dealing with the discretization problem of shadow maps, including percentage-closer filtering [7], which is one of the first methods introduced to alleviate this problem. This method is able to remove aliasing caused by shadow maps by applying a filter that averages the outcomes of depth comparisons against the shadow map. Percentage-closer filtering can be classified as a post-filtering approach since it applies the averaging after the non-linear depth test. One of the main drawbacks of this method is that the filtering operation can be executed only when the distance of the shaded point is available, so it should be repeated for every shaded point. Thus, the computational cost of percentage closer filtering becomes prohibitive when filters of large support are used. Unfortunately, relatively large kernels are needed in order to produce smooth anti-aliased shadows. To attack this performance limitation, we should move the filtering operation before the depth comparison and execute it once and globally for all shaded points. Such pre-filtering methods belong to two main branches, to those that apply depth transformation [8,9], or to those that are based on statistical analysis [10,11].

One of the most important representative of statistics-based shadow filtering is the method of variance shadow maps (VSM) [10]. VSM is based on using the Chebyshev’s Inequality for approximating an upper bound of the light visibility test. For each texel, VSM stores the depth and the squared depth of the shadow casters. These values can be filtered just like regular color textures resulting in the first two moments $M_1$ and $M_2$ of the depth values over the shadow filter region. When a point is shaded, the one-tailed versions of the Chebyshev’s Inequality allows us to upper bound the probability that depth $z_r$ of distribution with mean $\bar{z}$, and variance $\sigma^2$ is greater than the receiver depth $z_r$.

$$P(z_r \geq z_r) \leq \frac{\sigma^2}{\sigma^2 + (z_r - \bar{z})^2}.$$  

where $\bar{z} = M_1$ is the average of the depth values and $\sigma^2 = M_2 - M_1^2$ is their variance. If receiver depth $z_r$ is greater than the mean depth $\bar{z}$, then the variance shadow map method approximates the visibility function by this upper bound for probability $P(z_r \geq z_r)$

$$v_{VSM}(z_r) = \frac{\sigma^2}{\sigma^2 + (z_r - \bar{z})^2}.$$  

If the receiver depth is smaller than the mean depth, we assume that the surface is fully lit and thus $v(z_r) = 1$. VSM is an efficient and hardware friendly method. Its performance scales well with the screen resolution. Note that using an upperbound for the probability, the VSM may significantly overestimate the visibility function, making shadows lighter than they should be. This is noticeable as light leaks when shadow casters overlap from the light’s point of view. A typical problem occurs on parts of objects that are completely occluded but some amount of light still leaks inside the shadows (see Fig. 1).

Actually, light leaking artifacts show up because the first two moments of the depth do not provide enough information to disambiguate all the possible cases correctly. However, as stated in [12], storing more moments in the shadow map to obtain a sharper upperbound for the visibility function would not solve the problem, because higher-order moments are numerically unstable.

The layered variance shadow maps (LVSM) [12] method is an evolution of VSM developed to solve the “light leaking” problem. LVSM divides the light’s depth space into multiple layers, which allows for a better filtering of the two channels of the VSM. Using this technique we can obtain different upper bounds for $P(z_r \geq z_r)$, some tighter than others. When rendering the shadows, multiple
bounds allow for selecting the appropriate warp in which the light leaking artifacts are less visible or even eliminated. However, this technique introduces the problem of selecting the number and the optimal placement of the warps. Layers can be distributed by an automatic method based on the Lloyd relaxation algorithm. Although performance decreases and the storage cost increases as more layers are used, it can still perform the shadow filtering with just a single fetch from the texture memory. The LVSM paper also proposes to use an exponential warping technique instead of the layered approach called EVSM. The advantage of this technique is that it allows to remove the light leaking artifacts introduced by VSM while avoiding the complexity of the implementation and worse performance of the layered approach.

Convolution shadow maps (CSM) [9] linearize the problem by approximating the visibility function by a weighted sum of basis functions. One of the main advantages of this technique is that it supports hardware accelerated linear filtering, mip-mapping and anisotropic filtering. Instead of directly storing the depth values in the shadow map, a binary visibility function is encoded at each pixel. Then, these functions are approximated using a basis function expansion, which allows for the linearization of the shadow test. Summarizing, a depth transformation is applied to function expansion, which allows for the linearization of each pixel. Then, these functions are approximated using a basis in the shadow map, a binary visibility function is encoded at anisotropic filtering. Instead of directly storing the depth values supports hardware accelerated linear filtering, mip-mapping and approximating the visibility function by a weighted sum of basis functions into the map as in CSM or LVSM. Our method introduces a new statistical shadow map as in CSM or LVSM. Our method introduces a new statistical shadow map.

Exponential shadow maps (ESM) [8,13] approximate the visibility function with a single exponential. This technique is good for both shadow map pre-filtering and hardware accelerated filtering. Compared to convolution shadow maps, ESM is simpler and more efficient since it approximates the step function with just one term. Moreover, ESM is also more efficient than VSM and CSM in terms of memory footprint, as it only requires to store 1 single floating point value per texel, enabling for better performance because of fetching less amount of data. However, this technique still presents two major drawbacks: first, ESM is based on an assumption that is sometimes violated in certain cases (for example on non-planar receivers) which obligates the authors to fallback to a PCF approach; second, shadows artifacts appear at contacted shadows, which are diminishing as the distance between the shadow occluder and the shadow receiver increases.

Gumbau et al. [11] present a new statistical shadow map filtering technique that proposes a new approach for dealing with light leaking artifacts. It is based on approximating the cumulative distribution function of depths with a power function (see Fig. 2). This method stores the minimum, maximum and average values of the texels in the filtering area and uses these values to construct a power function that approximates the light visibility function. The visibility formula needs the computation of the cumulative distribution for the expected depth, which corresponds to

$$\hat{t} = \frac{z \hat{t} - z_{\min}}{z_{\max} - z_{\min}} + 1.$$

The visibility function obtained as

$$v_{\text{power}}(t) = \frac{1 - t^\beta}{1 - \frac{\beta}{\beta + 1}}$$

if \( t > \frac{\beta}{\beta + 1} \) and 1 otherwise, \( \beta \) is defined as

$$\beta = \frac{z - z_{\min}}{z_{\max} - z_{\min}} \leq 1.$$

Although this approach is able to significantly reduce light leaking artifacts with no additional computational or storage cost, it also has disadvantages. The use of the power function generates sharp anti-aliased shadows and it is difficult to control shadow smoothness.

The algorithm presented in this paper is a statistical method that allows for eliminating (or at least highly reducing in complex scenes) light leaking artifacts present in existing shadow map filtering techniques. In contrast to the previous work described in this section, our algorithm uses a Gaussian approximation which is able to approximate the depth distribution accurately and has small computational and storage cost. A similar approach was used in [14] to successfully approximate volumetric ambient occlusion for volumetric models, producing high quality smooth shading interactively. Compared to the approach based on the power function, our algorithm is able to generate smoother shadow boundaries.

3. Gaussian cumulative distribution

In order to obtain better shadows, our proposed method improves the evaluation of the stored statistical information to eliminate “light leaking” artifacts instead of putting more variables into the map as in CSM or LVSM. Our method introduces a visibility function approximation based on the VSM rather than using Chebyshev’s Inequality for upper-bounding the visibility factor. This approximation is based on a Gaussian distribution.
The application of probability theory techniques in shadow mapping is made possible by the observation that we can make a few fundamental assumptions on the unknown visibility function:

- At \( z = 0 \), that is when the shaded point is at the light source, the visibility function is 1.
- At \( z = \infty \), the visibility function can be assumed to be 0 since points at infinity are clipped away during rendering so it does not make any difference whether or not they are visible from the light source.
- The visibility function is monotonically decreasing, i.e. if the receiver is farther away, then there are more occluders between the light source and the receiver point, so the effect of the light source is possibly smaller, and is definitely not larger.

From the point of view of statistics, uncertainty is involved in the depth buffer, since the depth values are known only in depth map texel centers. The goal is to guess the visibility function at an arbitrary point by minimizing this inherent uncertainty. Unlike in signal processing, we are not constrained by linear operations, thus by the selection of a proper estimation, the depth testing can be made more robust.

In order to compute the visibility function, we rely on the probability of "no occlusion" \( P(z_0 \geq z_r) \). Unfortunately, this probability cannot be used directly since it would result in self shadowing due to the fact that for unoccluded planar surfaces, receiver depth \( z_r \) is equal to average depth \( \bar{z} \). So our distribution should give probability 1 for the mean. Note that a symmetric probability density would result in \( P(z_0 \geq z_r) = 0.5 \) and the probability would be close to 1 for extremely asymmetric distributions. Recall that classical shadow mapping cannot compare receiver depth to occluded depth since numerical precision issues and uncertainty of the depth buffer would cause incorrect decisions where the receiver and the occluder are on the same surface and are supposed to be exactly in the same distance, which generates shadow acne where no shadows are expected. The typical solution is involving some bias \( \beta \) in the comparison, which replaces uncertain cases by full visibility. Applying statistical filtering, the high frequency shadow acne is eliminated, but the uncertainty of shadow comparison is translated to the probability whether a point is in shadow. To fully avoid self shadowing, we should require \( v(z_0) = 1 \). Methods like VSM and ESM address the problem of self shadowing by defining the visibility function as a strong upper bound of \( P(z_0 \geq z_r) \).

Our visibility function should also be defined from the probability of "no occlusion" in a way that the self occlusions are eliminated. There are two ways of ensuring this, which can even be combined together: the application of conditional probabilities and biasing.

When \( z_r > \bar{z}_o \), random occluder distance \( z_o \) can either be smaller or larger than its average \( \bar{z}_o \). If \( z_o \leq \bar{z}_o \), then occlusion happens for sure. Assuming the other case, the visibility function can be expressed as a conditional probability provided that \( z_0 \geq \bar{z}_o \).

\[
v(z_r) = P(z_0 \geq z_r | z_0 \geq \bar{z}_o) = \frac{P(z_0 \geq z_r \cap z_0 \geq \bar{z}_o)}{P(z_0 \geq \bar{z}_o)},
\]

where we exploited that \( z_0 \geq \bar{z}_o \implies z_0 \geq \bar{z}_o \) when \( z_r > \bar{z}_o \).

If \( z_r < \bar{z}_o \), then \( v(z_r) = 1 \).

On the other hand, biasing would shift the value where the original visibility is evaluated. To make the biasing robust and adapt to the distribution of the depths, it is worth setting the bias proportional to the standard deviation \( \sigma \) of the random depth, i.e. we take \( v(z_r) = P(z_0 \geq z_r - \beta \sigma) \).

Both approaches have disadvantages. Simple biasing cannot guarantee \( v(\bar{z}_o) = 1 \) unless the probability distribution has a finite support. The conditional probability is better in this respect since it and avoids self occlusions, but the derivative of the visibility function changes abruptly at \( \bar{z}_o \) which produces sharper shadow boundaries. An aesthetic result can be obtained by combining the two approaches and biasing the conditional probabilities. If \( z_r > \bar{z}_o \), our proposed visibility approximation is

\[
v(z_r) = \frac{P(z_0 \geq z_r - \beta \sigma | z_0 \geq \bar{z}_o - \beta \sigma)}{P(z_0 \geq \bar{z}_o - \beta \sigma)},
\]

If \( z_r < \bar{z}_o \), then \( v(z_r) = 1 \).

Note that these arguments can be applied for any probability distribution \( P \). To obtain a specific implementation, we should select an appropriate distribution. Our selection is the Gaussian normal distribution since it has a separable probability density function that leads to efficient filtering, and can describe complex cases when many objects are independently distributed in space.

The probabilities are computed from the Gaussian cumulative distribution \( F(z; \bar{z}_o, \sigma^2) \) of random variable \( z_o \) having mean \( \bar{z}_o \) and variance \( \sigma^2 \).

\[
v_{\text{Gauss}}(z_r) = \frac{1 - F(z_r - \beta \sigma; \bar{z}_o, \sigma^2)}{1 - F(\bar{z}_o - \beta \sigma; \bar{z}_o, \sigma^2)} \text{ if } z_r > \bar{z}_o \text{ and } 1 \text{ otherwise.}
\]

The Gaussian distribution function can be expressed with the error function

\[
F(z; \bar{z}_o, \sigma^2) = \frac{1}{2} \left[ 1 + \text{erf} \left( \frac{z - \bar{z}_o}{\sqrt{2} \sigma} \right) \right],
\]

The error function, in turn, has the following approximation [15]:

\[
\text{erf}(x) \approx \frac{2}{\sqrt{\pi}} \frac{1 - \exp \left( -x^2 \right)}{1 + \exp \left( -x^2 \right)},
\]

where

\[
a = \frac{8(n - 3)}{3\pi(4 - n)} \approx 0.140012.
\]

Substituting the Gaussian CDF into our biased visibility formula we get

\[
v_{\text{Gauss}}(z_r) = \left[ 1 - \text{erf} \left( \frac{z_r - \bar{z}_o - \beta \sigma}{\sigma \sqrt{2}} \right) \right] / \left[ 1 + \text{erf} \left( \frac{\beta \sigma}{\sigma \sqrt{2}} \right) \right].
\]

Let us analyze the light leaking properties of this visibility function assuming the setting of Fig. 1, where objects A, B and C are in increasing distances \( a, b \) and \( c \) from the light source, B is partially occluded by A, and C is fully occluded by B. Thus, the shadow map contains only A and B, but not C. At the shadow boundary on object B, the average distance \( \bar{z}_o \) and the square root of the variance \( \sigma \) are

\[
\bar{z}_o = \frac{a + b}{2}, \quad \sigma = \frac{b - a}{2}.
\]

Substituting this into the visibility functions of classical variance shadow map and of the proposed one, and evaluating the visibility function for the distance of the fully occluded object, \( c \), we obtain

\[
v_{\text{VSM}}(c) = \frac{(b - a)^2}{(b - a)^2 + (2c - a - b)^2} = \frac{1}{1 + \left( \frac{c - b}{b - a} \right)^2}.
\]
4. Extending the Gaussian filter with warping

Although our Gaussian approximation is able to reduce light leaking artifacts with respect to variance shadow maps, there are scenes in which light leaking is still noticeable. These artifacts happen at those areas where several depth layers overlap (see Fig. 10).

This problem, which also occurs in the VSM approach, has been addressed by the authors of LVSM [12] by introducing a layered approach. However, to obtain the full potential of LVSM, we should solve the complex problems of how many layers are worth using and how to distribute the layers in depth space. The rendering performance decreases linearly as the number of layers increase, so the number of layers should be kept low. The optimal number of layers is unfortunately scene dependent, what is necessary for more complex scenes could be a performance killer for simpler virtual worlds, so it is a user parameter that must be set carefully for each scene.

Therefore, instead of incorporating a layered approach into our Gaussian approximation, we propose to follow the idea behind EVSM [12] and use a simpler warping function to further eliminate residual “light leaking” artifacts. The idea of warping is based on the fact that the visibility function is built of probabilities of “lower than” relations, i.e. \( P(z_o > z_r) \), and the relation remains the same if the involved variables are warped by a monotonically increasing function. However, when the probabilities are approximated, the transformation has an effect on the approximation. The idea is to still use the Gaussian CDF, but on values warped by an appropriate function \( w(z) \).

As concluded, the amount of light leaking depends on ratio \( r = (c - b)/(b - a) \), which should be as large as possible. If the average and the variance are computed for the warped values, and the Gaussian is also calculated for the warped receiver depth, then the ratio simply modifies as

\[
    r_{\text{warped}} = \frac{w(c) - w(b)}{w(b) - w(a)}.
\]

If the warping function is convex, then this ratio is increased with respect to \( r \), thus light leaking is reduced. The more convex the function is, the stronger the reduction of the light leaking artifacts on occluded object \( C \). Additionally, we wish to make shadow boundaries invariant to translations, so ratio \( r_{\text{warped}} \) must remain constant when all distances, \( a \), \( b \) and \( c \) are modified with the same

\[
    v_{\text{Gauss}}(c) = \frac{1 - \text{erf} \left( \frac{2c - a - b}{(b - a)\sqrt{2}} \cdot \frac{\beta}{\sqrt{2}} \right)}{1 + \text{erf} \left( \frac{\beta}{\sqrt{2}} \right)} \cdot \frac{1 - \text{erf} \left( \frac{2(c - b) + 1 - \beta}{(b - a)\sqrt{2}} \right)}{1 + \text{erf} \left( \frac{\beta}{\sqrt{2}} \right)}.
\]

In the ideal case, these functions should return zero if \( c > b \) since object \( C \) is in shadow, but they do not, which is responsible for light leaking. The visibility functions decrease with the increase of ratio \( r = (c - b)/(b - a) \) and the worst case happens when this ratio is zero. In the worst case the amount of light leaking is

\[
    v_{\text{VSM}}(c \approx b) = 0.5,
\]

\[
    v_{\text{Gauss}}(c \approx b) = \left[ 1 - \text{erf} \left( \frac{\beta}{\sqrt{2}} \right) \right] \left/ \left[ 1 + \text{erf} \left( \frac{\beta}{\sqrt{2}} \right) \right] \right.,
\]

which is in the range of \([0.32–0.44]\) when \( \beta \in [0, 0.5] \). Note that the amount of light leaking in the Gaussian model is smaller, but is not completely zero.

Fig. 3. Visual analysis of the curves used by the VSM and pure Gaussian approach for approximating the Heaviside step function introduced by the original shadow map technique. \( \sigma = 0.1 \).

Fig. 4. Visual quality comparison of our method using the spheres scene. Notice light leaking artifacts around the shadows of the spheres for VSM and how they are reduced with our algorithm, even though they are still noticeable. \( 5^2 \) filter kernel and \( 1024^2 \) shadow map sizes are used in the tests. (a) PCF. (b) VSM. (c) ESM. (d) Gaussian CDF.
value, which is satisfied by the exponential function. Therefore, an appropriate monotonic warping function is exponential $e^{cz}$ (where $c$ is a user-defined constant), as also suggested by [8].

However, if the warping function is too strongly convex, then shadowing artifacts show up on object B if it is a non-planar shadow receiver since the average or mean of the warped values will be significantly larger than the warped mean. The solution of this problem is the application of multiple warping functions similarly to [12,8]. We use the positive exponential $e^{cz}$ along with double negative exponential $-e^{-cz}$ because this has the effect of virtually moving object B towards receiver C, avoiding the non-planarity problem introduced by ESM. Fig. 9 shows the artifacts introduced by the exponential warping on curved surfaces and how it is solved using both the positive and negative terms.

Fig. 3 shows a visual analysis of VSM and our Gaussian methods where the function used for approximating the Heavy-side step function. A “good” visibility function should meet two contradiction requirements: it should be close to 1 and smoothly drop where the receiver depth is close to the average depth to produce smooth shadow boundaries. On the other hand,

![Fig. 5. Visual quality comparison on the chairs scene. Notice the artifacts introduced by VSM on the chair’s shadow. For ESM notice the loss of near occluders (highlighted) and the different shadow intensities of the chair and the table. $5^2$ filter kernel and 2048$^2$ shadow map sizes are used in the tests. (a) PCF. (b) VSM. (c) ESM. (d) Gaussian CDF.](image1)

![Fig. 6. Visual quality comparison on the car scene. Light leaking artifact introduced by ESM is highlighted with a red circle in (b). Our technique is able to remove light leaking artifacts on both VSM and ESM. $5^2$ filter kernel and 1024$^2$ shadow map sizes are used in the tests. (For interpretation of the references to color in this figure caption, the reader is referred to the web version of this article.)](image2)
it should converge to zero as fast as possible in order to eliminate light leaking, i.e. visibility significantly greater than 0 in regions that should otherwise be in shadow. The importance of the two contradicting requirements should depend on the statistics of the depth values nearby, i.e. the clever utilization of the variance $\sigma^2$ and possibly a user parameter (the bias in our case). The user parameter together with the variance determine how quickly and smoothly the visibility function is expected to fall from one to zero. Note that the curve produced by our method tends faster to zero than VSM, thus reducing its light leaking problems.

Fig. 7. Visual quality comparison of the different shadow map filtering techniques. The resolution for the shadow map is 1024 $\times$ 1024 for all tests. The filter kernel size is 5 $\times$ 5. Three different levels of zoom are provided for better examination of the results. (a) Percentage-closer filtering. Ground truth. (b) Variance shadow maps. Shadow leaking artifacts are introduced at shadow boundaries. (c) Exponential shadow maps. Shadow intensity artifacts introduced by differences on shadow map values. (d) Gaussian CDF. Improves the results shown in other techniques.
5. Implementation

In the implementation of the Gaussian method, first the pre-filtering stage performs a separable Box filter in two passes (horizontal and vertical) and averages the depth and the squared depth values in two separate channels of a texture map. In the case of the exponentially warped Gaussian filter, the depth is warped by functions $e^{cz}$ and $e^{-cz}$, then the same Box filtering is performed, but now resulting in a four channel map.

During shadow testing, we fetch the pair of the first and second moments of the depth and insert them with the receiver depth in the Gaussian formula (Eq. (3)). If warping is also used, the receiver depth is warped once with the positive exponential and once with the negative exponential, and we evaluate two Gaussian formulae using the positive and negative exponentially warped moments, respectively. The final value of the warped case is calculated as the minimum of the two probabilities.

6. Results

This section presents visual quality and performance tests for comparing our new filtering approach with well known existing techniques like variance shadow maps (VSM) and exponential shadow maps (ESM). The former uses the Chebyshev's Inequality approximation while the latter the exponential function for approximating the step function. Performance tests were generated on an Intel Core i7 Q9550 CPU @ 2.83 GHz with a NVIDIA GeForce 480GTX using Direct3D 10. The parameter $\beta$ is set to 0.5 in our formulation for all the tests. Visual quality tests are compared with PCF which is considered the ground-truth solution. In our PCF implementation, we perform distance-dependent filtering by taking into account the partial derivatives at a given screen pixel and then adapting the filter shape and size accordingly. Moreover, since our PCF implementation performs filtering using a fixed amount of samples per pixel, that would generate noise artifacts when filtering large areas of the shadow map. Therefore, in order to prevent these artifacts to appear, we have selected the appropriate test screens and kernel filter sizes so that the results of performing PCF generate smooth shadow boundaries.

Figs. 4-6 show a visual comparison of our approach with both VSM and ESM. It can be seen that the Gaussian CDF approach is more robust than both VSM and ESM, which exhibit artifacts in all the test scenes. Light leaking artifacts introduced by ESM in Fig. 6 are highlighted with a red circle. Note that our approach is not sensitive to light leaking artifacts introduced by multiple distant occluders as the exponential shadow maps (Fig. 7).

However, it must be noticed in Fig. 4(b) and (d) how our plain Gaussian method is not able of completely eliminating light leaking artifacts around the spheres, which motivated the development of the exponentially warped extension described in Section 4.

Fig. 10 shows a comparative visual analysis of our Exponential Gaussian approach compared to the reference PCF and EVSM. The results of the plain Gaussian approximation are also provided in order to show how light leaking artifacts are removed from the rendered scenes. This artifact is avoided in the prefiltering methods (VSM and our Gaussian approach) as anisotropic filtering can be used.

Our exponentially warped Gaussian approximation improves upon EVSM by the amount of reduction of light leaking artifacts. Fig. 8 shows different views of the Wires scene where EVSM is not able to completely eliminate light leaking artifacts while our technique is able to produce artifact-free shadows.

Table 1 compares performance data obtained with our approach (Gaussian CDF) with existing approaches: the well-known technique used in traditional percentage-closer filtering (PCF), the Chebyshev’s Inequality approximation used by variance shadow maps (VSM), the exponential approach used by exponential shadow maps (ESM) and,

![Fig. 8. Light leaking artifacts produced by EVSM. Our technique is able to generate artifact-free shadows under the same conditions. 92 filter kernel and 20482 shadow map sizes are used in the test. (a) Exponential Gaussian approximation. (b) Exponential Variance Shadow Maps.](image-url)
finally, the exponentially warped versions of VSM and our method. Different kernel-filter sizes as well as different shadow map resolutions have been used in the test.

All the times in Table 1 were measured using the scene shown in Fig. 10 at a screen resolution of 1920 × 1080. The reason of executing performance tests only for a single scene is that the performance of these methods is independent of the geometric complexity, and depends only on the screen and shadow map resolutions. Note that our Gaussian CDF approach is as fast as the traditional VSM approximation, because the amount of stored data on the shadow map is the same (two 32 bit values). The exponential shadow maps approach is always faster than other approaches because it only stores one 32 bit floating point value into the shadow map representing the exponential of the depth. This allows ESM to run faster by moving less data for each texture fetch operation than VSM and our approach which store twice that information.

The table shows that the Gaussian-based techniques clearly outperform PCF at low and medium shadow map resolutions. However, PCF performance does not depend on the shadow-map resolution, unlike the Gaussian filter which pre-filters the shadow map. Due to the higher computational cost of the \( e^x \) intrinsic function, the exponentially warped Gaussian filter runs just a little faster than PCF and is slower than the pure Gaussian method.

### 7. Conclusion

We have developed a new shadow filtering approach based on the statistical evaluation of variance shadow maps. Instead of upper-bounding with the Chebyshev's Inequality, we fit a Gaussian CDF on the depth distribution of the shadow map. This allows to enhance the visual quality of the shadows by reducing "light leaking" artifacts at no performance or storage cost penalty. Compared to exponential shadow maps, our technique has better visual quality at contact shadows but runs slower due to the amount of data to be transmitted in each texture fetch operation.

Note that the presented methods effectively (including VSM and ESM) perform anti-aliasing for the minification case (since we are able to use mip-mapping and anisotropic filtering) but not for the magnification case, where the soft shadow boundaries are achieved by blurring the shadow map using the different functions detailed in this article. For PCF, there is no true anti-aliasing done, since the blurring is done both for the minification and magnification cases.

Moreover, we have extended the idea of the Gaussian filter to further improve the visual results in complex scenes where the plain Gaussian method is not able to completely eliminate the light leaking artifacts. The extension we proposed is to combine our Gaussian filtering technique with exponential warps. We show how the combination of these techniques is able to further improve the results of the basic Gaussian approach improving the quality of the shadows on the rendered image. Comparing our exponentially warped Gaussian approach with EVSM, our exponentially warped Gaussian approximation improves upon EVSM in reducing the amount of light leaking artifacts. We show in Fig. 8 that EVSM is not able to completely eliminate light leaking artifacts of complex scenes while our technique effectively does.

In conclusion, the new technique presented in this paper produces good quality anti-aliased shadows at high frame rates, improves upon existing methods, and is scalable and reliable for complex scenes. These characteristics make it fully suitable for its use in real-time application such as games or VR.

As future work, we propose to attack the main drawback of our technique (and also of VSM and ESM), which is the performance bottleneck of shadow map pre-filtering. An interesting line of

### Table 1

Performance table of our Gaussian approaches compared to existing methods with different shadow map resolutions with varying kernel sizes. The scene shown in Fig. 10 was rendered in full-HD (1920 × 1080). All approaches used 32 bit floating point values for the shadow maps.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Filter size</th>
<th>256² (fps)</th>
<th>512² (fps)</th>
<th>1024² (fps)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PCF</td>
<td>3 × 3</td>
<td>932</td>
<td>921</td>
<td>908</td>
</tr>
<tr>
<td>Gaussian</td>
<td>3 × 3</td>
<td>760</td>
<td>691</td>
<td>557</td>
</tr>
<tr>
<td>VSM</td>
<td>3 × 3</td>
<td>770</td>
<td>704</td>
<td>580</td>
</tr>
<tr>
<td>ESM</td>
<td>3 × 3</td>
<td>802</td>
<td>725</td>
<td>635</td>
</tr>
<tr>
<td>Gaussian exp</td>
<td>3 × 3</td>
<td>728</td>
<td>600</td>
<td>475</td>
</tr>
<tr>
<td>EVSM</td>
<td>3 × 3</td>
<td>755</td>
<td>682</td>
<td>502</td>
</tr>
<tr>
<td>PCF</td>
<td>9 × 9</td>
<td>307</td>
<td>311</td>
<td>302</td>
</tr>
<tr>
<td>Gaussian</td>
<td>9 × 9</td>
<td>712</td>
<td>621</td>
<td>481</td>
</tr>
<tr>
<td>VSM</td>
<td>9 × 9</td>
<td>731</td>
<td>656</td>
<td>502</td>
</tr>
<tr>
<td>ESM</td>
<td>9 × 9</td>
<td>782</td>
<td>682</td>
<td>612</td>
</tr>
<tr>
<td>Gaussian exp</td>
<td>9 × 9</td>
<td>700</td>
<td>575</td>
<td>398</td>
</tr>
<tr>
<td>EVSM</td>
<td>9 × 9</td>
<td>728</td>
<td>605</td>
<td>413</td>
</tr>
<tr>
<td>PCF</td>
<td>15 × 15</td>
<td>128</td>
<td>127</td>
<td>124</td>
</tr>
<tr>
<td>Gaussian</td>
<td>15 × 15</td>
<td>703</td>
<td>608</td>
<td>456</td>
</tr>
<tr>
<td>VSM</td>
<td>15 × 15</td>
<td>715</td>
<td>612</td>
<td>478</td>
</tr>
<tr>
<td>ESM</td>
<td>15 × 15</td>
<td>733</td>
<td>631</td>
<td>507</td>
</tr>
<tr>
<td>Gaussian exp</td>
<td>15 × 15</td>
<td>685</td>
<td>497</td>
<td>300</td>
</tr>
<tr>
<td>EVSM</td>
<td>15 × 15</td>
<td>713</td>
<td>503</td>
<td>342</td>
</tr>
</tbody>
</table>

![Fig. 9](image) Taking into account both the positive and negative exponential terms allows our technique to render properly anti-aliased shadows on both planar and non-planar receivers. Notice the shadowing artifacts in non-planar surfaces when using only the positive exponential term \( e^x \). 7² filter kernel and 1024² shadow map sizes are used in the tests. (a) Exponential Gaussian (only positive), (b) Exponential Gaussian (positive and negative).
research would be to study how to diminish the penalty of larger depth map resolutions. Also, there is still room for improving the current method by reducing the amount of information needed in the shadow map. Moving from the pure Gaussian filtering to the exponentially warped version forces us (and EVSM) to duplicate the amount of calculations performed: the positive and negative

Fig. 10. Comparison of the results obtained by combining both the Gaussian filtering and the exponential warp function on two different scenes. The shadows calculated with PCF as well as with EVSM are also provided. Software built using a base implementation from [16]. 9° filter kernel and 2048° shadow map sizes are used in the test. (a) Pure Gaussian filtering. Visible shadow leaking artifacts are introduced at shadow boundaries. (b) Exponential Variance Shadow Maps. Reduces the amount of light leaking artifacts of pure Gaussian filtering. (c) Exponential Gaussian filtering. Further improves the results of EVSM. (d) Percentage-Closer filtering (PCF) provided as "ground truth" for visual comparison.
contributions of the exponentially warped two moments. Finding a way to avoid processing two times the input data would increase the overall performance of the method. Finally, we believe it is worth studying other statistical approaches in order to maximize performance while minimizing artifacts on the rendered image.

Acknowledgements

This work has been supported by the Spanish Ministry of Education and Science (TIN2010-21089-C03-01, TIN2010-21089-C03-03, TIN2009-14103-C03-01), Caja Castellón-Bancaja Foundation (P1.1B2010-08, P1.1B2009-34), the European Union (Ref. 226487), Generalitat Valenciana (Project PROMETEO/2010/028, BEST/2011/287, Generalitat of Catalunya (Catalan Government) (2009-SGR-643), and TAMOP - 4.2.2.B-10/1-2010-0009.

References


