Realistic Water Volumes in Real-Time
Lionel Baboud, Xavier Décoret

To cite this version:
Realistic Water Volumes in Real-Time

Lionel Baboud and Xavier Décoret
ARTIS-GRAVIR/IMAG-INRIA

Figure 1: Examples of rendering of water volumes. All images are 800 × 600 and are generated at about 30Hz.

Abstract

We present a real-time technique to render realistic water volumes. Water volumes are represented as the space enclosed between a ground heightfield and an animable water surface heightfield. This representation allows the application of recent GPU-based heightfield rendering algorithms. Our method is a simplified raytracing approach which correctly handles reflections and refractions and allows us to render complex effects such as light absorption, refracted shadows and refracted caustics. It runs at high framerates by exploiting the power of the latest graphic cards, and could be used in real-time applications like video games, or interactive simulation.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism

1. Introduction

Realistic rendering of water is one of those effects that can highly increase the perceived realism of virtual scenes. However, the appearance of water is caused by many physical phenomena that were typically not affordable in real-time until recently. Indeed, the increasing functionalities and power offered by graphics hardware allow the adaptation of well-known off-line techniques to real-time applications. In particular, ray-tracing is well adapted to the SIMD nature of fragment processors. However, porting generic ray-tracers to the GPU is quite involved [PBMH02]. Hierarchical acceleration data structures, or generic mesh-ray intersection routines, must be carefully rewritten to fit the GPU organisation. Fortunately, in specific situations, the computations can be greatly simplified and efficiently mapped to the GPU. In particular, this is the case for the ray-tracing of heightfields.

The question is then to identify the class of objects that can be adequately represented with these simpler models.

The contribution of this paper is to show that water basins and similar structures can be quite realistically modelled by two heightfields, one for the ground, and one for the water surface. Using adaptations of recent techniques for heightfield rendering, complex effects like reflections, refractions, light absorption, refracted shadows and caustics can be efficiently simulated in real-time. Except for a part of the caustics computations, it runs entirely on the GPU. The amount of displayed geometry is negligible and the rendering time is proportional to the number of pixels occupied by the water on screen. The different features are mostly independent and can be enabled separately, allowing a tradeoff between realism and performance.
2. Previous work

Displaying water surfaces involves two orthogonal problems: simulating the movement and rendering it accurately. Our work focuses only on this latter part, and any animation method can be used. We briefly review existing techniques.

Water Simulation Fluids are governed by the Navier-Stokes equations which have been used at different levels of simplification. Applying these equations to a volumetric representation of water leads to physically accurate but computationally intensive simulations [FF01, EMF02]. They are required for complex movements like breaking waves or splashing water. The study of Navier-Stokes equations in the context of ocean surfaces modeling has led to the Gerstner and Biesel swell model which decomposes ocean waves into a sum of trochoidal wave trains. Realistic results [FR86, HNC02] have been obtained by combining it with models expressing waves amplitude, frequency and direction spectra derived from measurements [PM64, HDE80], and waves propagation properties like refraction near coasts [Pea86, TB87, Tes99, GM02]. These approaches generally allow one to compute a heightfield representation of the water surface in real-time. In the case of bounded water surfaces, the surface heightfield can be expressed as the solution of simple partial differential equations [KM90] which can be solved numerically by efficient local diffusion schemes [Gom00, Sta03]. The spectral approaches [MWM87, Tes99, PA00] first compute a Fourier spectrum of the water surface heightfield, using physically measured distributions, before transforming it into a heightfield with a Fast Fourier Transform. The resulting heightfield is bounded but periodic.

Water Rendering Rendering of water surfaces is difficult due to complex interactions with light. As explained by [PA00], the physical study of light-water interactions is a full-fledged research field with a vast literature. Several works have focused on the transcience of these phenomena for computer generated images [Wat90, PA00, IDN03]. Although realistic, these approaches are computationally intensive. Realtime solutions exist, but they generally make important simplifying assumptions [NN94, HVT’04]. Reflection and refraction are important effects for the realism of virtual water. Until recently, reflection approaches would only address special configurations such as planar reflectors [FvdDH90] or infinitely far reflected objects, or would be only interactive [HNC02] or based on precomputations [HLCS99, YYM05]. Recently, real-time approaches using GPU-based ray-tracing and image based approximations of the reflected objects have been proposed [SKALP05, PMDS06]. Refraction algorithms usually consider one transparent object surrounded by an infinitely far environment represented by an environment map (see [GLD06] for a recent review of these methods). This assumption cannot be used in the case of the refraction of a ground surface through a water volume. Computing the refraction of nearby objects can be done with memory costly light-field precomputations [HLCS99]. Caustics are caused by the convergence of light due to reflections and refractions. They are computationally expensive to render and are traditionally handled by global illumination algorithms like photon mapping [Jen01]. Interactive solutions exist based on different photon emitting and gathering strategies [EAMJ05, WS03, WD06, DS06] but they need to reduce the number of emitted photons to run in realtime, thereby producing noisy or blurred caustics. We use a method conceptually very similar to backward beam tracing with illumination maps [Arv86].

Heightfield rendering Our technique is based on efficient heightfield rendering. Efficient GPU-based implementations [POC05, Tat06, BD06] have made it an advantageous way to render geometric details with small computation and memory costs. These algorithms sample the heightfield texture along each viewing ray using different sampling strategies. [POC05] intersect the viewing ray with fixed horizontal planes before computing a precise intersection with few iterations of a binary search. This amounts to an even sampling in the vertical direction, which can cause stair-stepping artifacts at grazing angles. This dependence of the sampling on the view angle is reduced by [Tat06] which chooses the number of samples between fixed bounds according to the viewing ray direction. The horizontal sampling method described by [BD06] uses precomputed information to sample the ray optimally. This allows the accurate display of heightfields potentially containing fine features.

3. Our approach

3.1. Modelling hypothesis

Our observation is that, except for big waves, water surfaces are mostly 2.5D. Similarly, because of gravity, most terrain can be represented as a displacement over a flat surface. Therefore, we choose to represent a water volume with two heightfields, the ground surface and the water surface, both encoded as 2D textures (Fig. 2). The texture for the water surface is the output of a simulation procedure which can be procedural (e.g. a sum of basis functions) or numerical (e.g. a physical simulation). To render the water volume, we render its bounding box and use a fragment shader to perform
ray-tracing. During this, we only consider one level of recursivity: for each viewing ray, we find the intersection with the water surface, then the intersection of the refracted ray with the ground. Avoiding recursion is important because current fragment shaders do not support recursive function calls. This single level recursion assumption is correct if the ground is non-reflective, and if no light-ray crosses the water surface more than once. Due to the leaning effect induced by refraction, even in presence of big waves, this hypothesis generally holds (Fig.3). When it is not the case, the visual difference induced by this simplification is hard to notice for the human eye. The rays we consider are shown on Figure 4. The algorithm outline is the following:

```
for each screen pixel
    d ← direction of the corresponding viewing ray
    ps ← intersection (ground, ray (v, d))
    pn ← intersection (water, ray (v, d))
    if (ps undefined and pn undefined)
        do nothing (discard fragment)
    else if (ps before pn)
        pixel ← lightedGroundColor (ps, d)
    else
        dr ← refractedDirection (d, n (ps), η)
        di ← reflectedDirection (d, n (pn))
        ps ← intersection (ground, ray (pn, di))
        pn ← intersection (ground, ray (pn, dr))
        C1 ← lightedGroundColor (ps, di)
        C2 ← if (pn undefined)
            envMap (dn)
        else
            lightedGroundColor (pn, dr)
        F ← fresnelReflectivity (d, n (ps), η)
        pixel ← (1 − F) × C1 + F × C2
```

The fresnelReflectivity procedure computes the ratio $F$ between reflected and incident radiance with the Fresnel formula. For a fixed refraction index $η$, it is only a function of the incident angle, which can be efficiently computed with the approximation given by [Sch93] or simply sampled in a precomputed 1D texture. The lightedGroundColor procedure computes the color of a point on the ground from its coordinates, the viewing direction and the light position. True computations should integrate all rays coming from the light and arriving at $p_n$ after refractions. Instead, we make a simplification. We use a simple Phong model involving the ray from $p_n$ to the light, the normal, and the material at that point (the latter two being specified by texture maps). Only the amount of light is computed, taking into account the many possible light rays (see Section 4.3 on caustics).

Note that we implicitly supposed that each ray entering the water hits the ground further on. This condition is satisfied if the water is correctly enclosed by the ground surface, i.e. if on the borders the ground height values are higher than the water ones. This holds for a basin or a pond, but not for an aquarium (Fig.5). In that latter case, a dynamic environment map can be used to approximate what the ray “sees” when exiting the water.

3.2. Optimized rendering

Our rendering algorithm is a tailored version of [BD06]. We refer the viewer to that paper for detailed description. In brief, it walks along the ray until a binary search can be safely run, i.e. it is guaranteed that the interval of that search contains at most one intersection with the heightfield. For that, it uses a pre-computed safety radius texture indicating for each texel the maximum neighborhood in which a ray
above the texel can intersect the heightfield no more than once. The paper also describes a slower – but still exact – method that does not use the safety radius texture.

In the rendering of water surface, we perform intersections with two heightfields, the ground and the water surface. As the ground is static, precomputing the safety texture can be done off-line. This is not the case for the water texture which is dynamic. Fortunately, compared to the ground, it generally has a small amplitude. We thus estimate the average amplitude and size of the features, and compute a constant safety radius accordingly. In the worst case, it gives a radius of 1, which boils down to using the exact intersection described in [BD06]. To further optimize the number of steps required to find the intersection, we intersect the viewing ray with the bounding box of the water surface instead of that of the ground and water altogether.

Conceptually, we run the ray intersection algorithm several times (see pseudo-code in Section 3). However, in practice, we can perform the search along the ray for the two heightfields at the same time, thereby saving the computations of the successive samples taken on the light ray. It allows to stop as soon as the first of the two surfaces is hit. Moreover, if the two heightfields are packed in a single texture, this divides by two the number of texture lookups performed. Packing the ground heights with the water ones can be done during the precomputing of the water texture at no additional cost.

We want to emphasize that various strategies can be adopted here. Although we use the artifact-free approach of [BD06], any other method such as [POC05, Tat06] can be used. One can also choose to use different methods (binary search, horizontal or vertical fixed steps, Amanatides traversal, constant or texture-based safety radius, etc.) depending on the desired tradeoff between speed and quality, and depending on possible constraints on particular viewing conditions (distant or nearby water, grazing angles, large or small amplitudes of the heightfields, etc.). The only limitation is that the joint intersection described in previous paragraph is applicable only if the same method is used for intersecting with the ground and the water. Our goal in this paper is to show that the availability of these techniques and the heightfield representation for water basin and alike makes the rendering of complex effects achievable in real-time. We provide images and timings of what the programmers can expect, but the tuning for a particular application remains their responsibility.

Finally, we would like to say a few words about the precision used for heightfield values. For most situations, using one 8 bit channel per heightfield is sufficient. But if the ground or water surfaces are to be viewed from very close, or if they have large amplitudes, it will not be sufficient. Even if the ground is smooth and is correctly sampled, the bilinear interpolations performed on 8 bits numbers will produce stairs artifacts. This can be avoided by using 16 bits textures (the two height maps can be packed in a 32 bits texture with two 16 bits channels). Note also that as water is likely to have a small amplitude, it is more appropriate to store it in a normalized form by passing two additional parameters to the shaders : a height offset (the mean water level) and an amplitude. Doing so allows us to fully use the precision offered by the texture.

3.3. Integration with other objects

For our method to be integrated within a classically modelled scene, we must handle correctly the interactions with other objects. The first thing to notice is that for each fragment, the exact 3D intersection point of the viewing ray with the first encountered surface is known. Transforming it into the camera frame, we can trivially compute an appropriate z-value (instead of that corresponding to the bounding box we are actually drawing). Furthermore, for rays that do not intersect any of the two surfaces, the corresponding fragment is discarded, so it does not interfere with any past or future value in the framebuffer. Finally, the current depth in the z-buffer can easily be taken into account during the intersection procedure. We just stop the walk along the ray when this value is exceeded. For all these reasons, our water volumes seamlessly integrate with the z-buffer, yielding correct occlusions with other rendered primitives.

This statement is however true only for objects that do not enter totally (like fishes) or partially (like the legs of a character) the water volume. Treating these cases is difficult because after undergoing the refraction, the viewing rays can be scrambled in a way which is difficult to predict (Fig.6). Having objects occluding viewing rays between the water surface and the ground breaks the organization into heightfields of this specific volume which made the raytracing feasible. Note that this effect is typically faked with ad-hoc hacks in games.

![Figure 6: Objects in water are difficult to correctly display because of refraction.](image-url)

4. Simulated effects

We now discuss a list of effects that can be straightforwardly simulated with our approach. Each effect can be turned on or off in the shaders, allowing for a tradeoff between realism and speed.
4.1. Underwater light absorption

Careful attention has to be paid to lighting phenomena occurring in water. As explained in [PA00], some of these phenomena, like light scattering, are difficult to handle. Thus lighting equations have to be simplified. We use the lighting model given by [PA00] for a given wavelength \( \lambda \), the radiance transmitted from a point \( p \) under water to a point \( p_w \) on the surface is given by:

\[
L_\lambda(p_w, \vec{0}) = \alpha_\lambda(d, 0)L_\lambda(p, \vec{0}) + (1 - \alpha_\lambda(d, z))L_{d\lambda}
\]

where \( \vec{0} \) is the direction from \( p \) to \( p_w \), \( L_\lambda(p, \vec{0}) \) is the outgoing radiance from the point \( p \) in the direction \( \vec{0} \), \( d \) and \( z \) are respectively the distance and the depth difference between \( p \) and \( p_w \). \( L_{d\lambda} \) is a constant diffuse radiance computed from scattering measurements involving sky and sun colors, and \( \alpha_\lambda(d, z) \) describes an exponential attenuation depending on traveled distance and depth:

\[
\alpha_\lambda(d, z) = e^{-a_\lambda d - b_\lambda z}
\]

where \( a_\lambda \) and \( b_\lambda \) are attenuations coefficients depending on water properties. The first term of the equation equates radiance coming from \( p \) attenuated with traveled distance, the second one equates the contribution of diffuse scattering, depending also on depth.

Remember that during the rendering process we compute the precise coordinates of \( p \) and \( p_w \) so traveled depth and distance are easy to compute. This makes the implementation of this lighting equation costless relative to other computations.

Light attenuation can vary with wavelength, producing visually compelling chromatic scales. Ideally the color spectrum has to be discretized and computations must be done per wavelength. Using only the three components of the RGB color space gives an approximate approximation. This is done by computing the attenuation color:

\[
\alpha(d, z) = (\alpha_G(d, z), \alpha_G(d, z), \alpha_B(d, z))
\]

and combining it component-wise with the incoming and diffuse colors. Precisely desired color variations (e.g. physically measured) can be obtained by using a precomputed 2D color texture containing sampled values for \( \alpha(d, z) \).

In practice, the influence of depth (the \( b_\lambda z \) term) is subtle and considering only traveled distance can be sufficient. Doing so simplifies the attenuation coefficient to:

\[
\alpha_\lambda(d) = e^{-a_\lambda d}
\]

which can now be sampled in a 1D color texture (Fig.7). With this simplification, \( L(p_w, \vec{0}) \) is the simple linear blending between \( L(p, \vec{0}) \) and \( L_{d\lambda} \) with respect to \( \alpha(d) \).

Notice that \( a_\lambda \) can be seen as a distance scaling factor, so it can be scaled according to the scale of the water volume. If this value is large, it means that light will be quickly absorbed so the deep parts of the ground surface will not be visible. In that case it is useless to keep running on the refracted ray after a certain limit distance \( d_{\text{max}} \). This distance can be defined by fixing an attenuation threshold \( \varepsilon \) under which the incoming ground color is considered to have a negligible contribution:

\[
\forall \lambda, \quad e^{-a_\lambda d_{\text{max}}} < \varepsilon
\]

hence simply computable as:

\[
d_{\text{max}} = \max \left( \frac{-\log(\varepsilon)}{a_\lambda}, \frac{\log(\varepsilon)}{\min a_\lambda} \right)
\]

4.2. Self-reflections

Reflections on the water surface can be separated into three classes: reflections of distant objects, reflections of near objects and reflections of the parts of the ground emerging from water (Fig. 8). The first class can be easily handled using an environment map whereas only heuristics exist for the second one. We call the third class self-reflection. This case is more specific: emerging parts of the ground are represented in the ground heightfield, which allows us to perform a rapid raytracing of the reflected ray, as we do for refracted rays. So to take these reflections into account, the reflected ray is intersected with the ground surface. If an intersection exists, the corresponding lighted ground color replaces the one that would be otherwise looked up in the environment map.

\[
\text{Figure 8: Reflected ray hitting an emerging relief.}
\]
4.3. Caustics and shadows

Computing caustics is hard to carry out in a forward ray-tracing approach because it requires us to count the amount of incoming light at each rendered point. That is why most existing caustics algorithms are based on backward raytracing: rays are cast from the light source instead of the viewpoint. We use a two-pass photon-mapping-like algorithm, involving GPU/CPU transfers of textures. Fortunately, we only need small textures and this does not impede the performance too much. The first pass renders the water surface from the light source into a photon texture. The texels of this texture record the coordinates of where the corresponding light rays hits the ground. Since the ground is a heightfield, we need only recording the \((x, y)\) coordinates, which leaves the third channel of the texture available to store the photon contribution, based on the Fresnel transmittance, the traveled distance, the incident angle and a ray sampling weight. We apply a random jittering to light rays to reduce aliasing artifacts (see Figure 9). We then gather the photons recorded in the texture. This boils down to computing a two-dimensional histogram of this texture. Unfortunately, it can not be done efficiently on the GPU because of the lack of forward mapping, so the photon texture is transferred to CPU where it is processed to construct an illumination texture (by analogy to illumination maps of [Arv86]). We traverse the texels of the photon texture, retrieve the position and intensity of the corresponding photon, and add this intensity to that stored at that position in the illumination texture. The illumination texture is then transferred to the GPU where it is used for lighting in the final render pass.

The illumination texture can be very noisy due to the crude sampling of our approach. Applying a simple gaussian blur would remove noise but would also blur caustics patterns which, by nature, present high frequencies. That is why edge preserving filters like anisotropic diffusion or bilateral filtering [TM98, DD02, WMM*04] are traditionally used to filter density maps in photon mapping techniques. We apply a GPU-based bilateral filter which gives good results once correctly tuned. The whole process is summarized in Figure 10.

![Figure 9: Photons are emitted from jittered pixel positions and mapped on the ground surface.](image)

The resolutions for the two intermediate textures must be adjusted to balance the desired quality and framerate. The resolution of the photon texture directly determines the number of photons. The more photons are used, the more accurate the caustics will be. The illumination map can be seen as a lighting texture directly mapped on the ground, so its resolution affects the fineness of the reproduced caustic patterns. Note however that these two resolutions are linked: if too few photons are emitted, a large caustics map will be noisy. The resolutions are limited by the cost of their transfer between GPU and CPU memory. Using a 256 \times 256 photon texture and a 128 \times 128 illumination texture gives good results in practice.

This approach for computing caustics also accounts for shadows cast by the ground on itself, with refractions handled. Shadows of other objects in the scene onto the ground can not be cast that easily. Once again, the difficulty is to handle the refraction of rays. However, one could use a shadow map approach as a coarse approximation. It would integrate directly with our algorithm by performing a projective texture lookup into the shadow map, using the world position where the ray hits the ground.

5. Results and discussion

We implemented the proposed algorithm using GLSL shading language, and ran the example on a GeForce 7800FX. We implemented animation using [Gom00] with forces...
specified by random rain-drops or by mouse interactions (see accompanying videos). Table 1 indicates performance measurements. The cost is highly dependent on the number of fragments rasterized. For our tests, the displayed volume covers a large part of a 800 × 600 window, yet it renders in real-time. Another interest of the method is that its memory requirements are low: only a couple of textures are required. In all examples shown, we used 128 × 128 textures for the ground and water surfaces, yet they look appealing.

<table>
<thead>
<tr>
<th>coverage</th>
<th>refraction</th>
<th>+ auto-reflection</th>
<th>+ caustics</th>
</tr>
</thead>
<tbody>
<tr>
<td>100%</td>
<td>26 Hz</td>
<td>20 Hz</td>
<td>15 Hz</td>
</tr>
<tr>
<td>50%</td>
<td>45 Hz</td>
<td>36 Hz</td>
<td>23 Hz</td>
</tr>
<tr>
<td>25%</td>
<td>87 Hz</td>
<td>69 Hz</td>
<td>34 Hz</td>
</tr>
</tbody>
</table>

Table 1: Framerates obtained at 800 × 600 for different screen coverages and with various effects turned on successively (rightmost column combines all effects).

In all examples shown, we used 128 × 128 textures for the ground and water surfaces, yet they look appealing.

Figure 1 and Figure 11 show the kind of effects we can simulate. We believe the accompanying videos should also convince that a high level of realism is achieved with this technique. In particular, relevant effects are simulated, and the illusion of water is convincing.

6. Conclusion and future work

In this paper, we show how the encoding of water volumes with two heightfields allows for real-time rendering of realistic water. Using an efficient ray-tracing approach for such a representation, we show that many effects can be combined: single bound refractions and reflections, fresnel effect, light absorption, caustics and shadows. None of these effects are new; the contribution of this paper is to validate the amenability and suitability of heightfields in this context. Each effect can be turned on or off depending on the desired tradeoff between speed and quality. The method is based on simplifications of the underlying physical phenomena, yet gives very appealing results. The rendering cost is directly proportional to the screen occupancy of the displayed water. Thus, we believe it can be of great interest for rendering elements of a virtual world such as aquariums, ponds, muddy puddles, fountains, etc. in a much more realistic way than previously. Yet, even at large resolutions, it is still fast and may also prove useful for sea and lake rendering.

In the future, we would like to investigate a couple of optimizations (we plan to release the shaders and a detailed description of the formulae used). We also want to generalize the approach to other objects that combine a transparent layer in front of an opaque one, such as windows in a city walkthrough. We also want to tackle the problem of geometry-based objects that penetrate the water, for which we investigate approaches that dynamically “project” the geometry onto the heightfields.

References


