Selective Tone Mapper

Author: Alessandro Artusi, Intercollege Cyprus, artusi.a@intercollege.ac.cy

Author: Benjamin Roch

Author: Yiorgos Chrysanthou, University of Cyprus, Yiorgos@ucy.ac.cy

Author: Despina Michael University of Cyprus,

Author: Alan Chalmers, University of Bristol, alan@compsci.bristol.ac.uk

Version: 1.1-English
Date: 31.03.2007

Index

Abstract	**p.**
Introduction	**p.**
Related Work	**p.**
Selective Tone Mapper	**p.**
Framework	**p.**
Experimental Results	**p.**
Conclusions and possible future	**p.**
Bibliography	**p.**

Abstract

Recently human visual attention has been increasingly exploited in computer graphics in order to reduce the high cost of computing high-fidelity images. By only computing in high quality those areas of a scene which are perceptually important, significant computation time can be saved without the user being aware of the quality differences within the image. Tone mapping (TM) is a key part of the high-fidelity image synthesis process, allowing high dynamic range images to be best displayed on low dynamic range computer monitors. Although it is possible to achieve, by means of graphics hardware support, interactive tone mapping using lower quality global tone mappers, to-date interactive performance for the superior quality, but computationally significantly more expensive, local tone mapping has been possible but only at a much reduced quality. This report presents the novel concept of Selective Tone Mapper, that it is a general framework, for reducing the computational costs of the existing local TM operators. We propose a GPU implementation of this framework which overcomes many of the problems modern graphics hardware have, in order to fully exploit the advantages of the Selective Tone Mapper concept.

Introduction

Visual Attention is the ability of the Human Visual System (HVS) to find and focus on relevant information quickly and efficiently [Re02]. The most important aspect of the visual attention is its capacity to rapidly select objects of interest in our visual environment. Objects which are not perceptually important can often go completely unnoticed [MR98]. Visual attention models have been used to significantly reduce the overall computation time for global illumination calculations without a perceptible loss of image quality, for example [Yee+01][Ca+03]. Tone mapping operators (TMOs) are a key part of the process of high-fidelity image synthesis, as they attempt to generate images visually similar to a real scene by careful mapping to a set of luminances that can be displayed on a low contrast ratio display or printed. TMOs may be classified as *global* or *loca*l. Global TMOs apply the same operation to all pixels of the input scene, while local operators take into consideration the local properties of individual pixels and use this information to preserve the local contrast reproduction. Although with graphics hardware support it is indeed possible to achieve interactive tone mapping using global operators, for example [Ar+03], the computational expense of local TMOs have so far limited their use in interactive applications. These limitations are directly connected with the limitations of the current graphics hardware. In fact a straight forward hardware implementation of the original CPU implementation of the TMOs is not possible, and often significant modifications are required in order to achieve the hardware implementation. This results in a much reduced quality when compared with the original TMO, and the time performance is rapidly decreasing as the resolution of the input frame increases. In this paper we introduce the novel concept of a Selective Tone Mapper which uses aspects of visual attention models to direct local TMOs at the perceptually important parts of a scene while a global TMO is used for the remainder.

This idea results in a general framework that may be applied to any pre-existing local TMOs. We will show how to make use of the current graphics hardware in order to fully implement the framework on the GPU.

In this way we are able to achieve perceptually high quality tone mapping at real-time rates.

To this work there is a Patent pending GB 0709392.5, Title: Selective Tone Mapping

Related Work

Visual attention is the process by which the HVS rapidly finds and focuses on perceptually important information in a potentially cluttered environment [Re02]. Our eyes do not scan a scene in a raster-like fashion, but rather they jump so that important parts of the scene fall sequentially onto the fovea, the area of highest visual acuity within our eyes. There are two distinct visual attention processes, which [Ja90] termed bottom-up and top-down. The bottom-up process is purely stimulus driven. This is evolutionary, enabling our eyes to rapidly identify salient information in a scene without volitional control.

The top-down process is, on the other hand, under voluntary control and enables our visual system to focus on parts of a scene which may be relevant to accomplish a task within that scene.

Inattentional blindness is a feature of the HVS which may result in objects in a scene which would normally attract attention, being completely ignored if they are not relevant to the current task [MR98][SC99].

In computer graphics, early work, including the *Visual Difference Predictor* [Dal93][My98] and the *Visible Discrimination Model* [BM98], used knowledge of the HVS to investigate the perceptual quality of images.

More recently, researchers have used detailed models of the bottom-up (known as Saliency Maps [IKN98]) and top-down (known as Task Maps [Ca+03]) visual attention processes to significantly reduce the computation of high-fidelity global illumination calculations.

Cater et al. [Ca+03] introduced the concept of a task map which exploit inattentional blindness, Sundstedt et al. [Su+04] were the first to combine both saliency and task maps. Both of them make use of this information to reduce the overall computation time in image rendering.

In addition, Longhurst et al. [long06] developed an hardware implementation of the importance map for interactive applications.

It is not the purpose of this paper to give a complete overview of the state-of-art of the TMOs proposed in the literature. For a full overview of tone mapping see [devlin2002].

In this Section we will concentrate on reviewing the work that attempts to develop a real time TMO and in particular a local TMO.

The local TM methods are concerned with achieving some form of perceptual accuracy of individual visual effects. This requires significant computational effort and none of these TMOs operate at interactive rates. Some authors, including [GWW*03][kra+05] implemented local TMOs tightly coupled with current graphics hardware achieving interactive rate only for very low frame resolution.

Recently Roch et al. [Ro+07] presented an implementation of a state of art local operator [Ash02] that attempts to reduce its computation time using an efficient GPU implementation of the Gaussian filter used for the computation of the local luminance adaptation.

Cadik [Ca07] presented an approach that validate the main idea of visual attention in the context of TM application. Our approach differs from the one presented by Cadik since we analyse how to have a general framework, applicable to any local pre-existing GPU TMOs, tightly coupled with current graphics hardware.
We integrated it in an interactive system, and we solved a number of issues that arose.
In detail we analysed how to overcome the problems of current graphics hardware in order to have an efficient GPU implementation of our framework, demonstrating the ability to accelerate a GPU implementation of a pre-existing local TMO.

Selective Tone Mapper

In general, pre-existing local TMOs also have a global component. The local component is the gain control; the global one uses this gain in order to compress the high dynamic range of the input scene into the low dynamic range available in the display. In our Framework we apply the costly computational component of the local TMOs only on the parts of the input scene that attract the attention of the HVS as identified by the local areas localization algorithm.
The global component of the TMOs is applied to the remainder of the scene which are below the ``high perceptual threshold".
This framework has the following significant advantages: first, to be independent from the TMO used, second the computational time is strongly reduced due to the fact that the high computational cost is on a reduced number of pixels. Finally it can be integrated in the rendering pipeline without any modification of the pipeline.

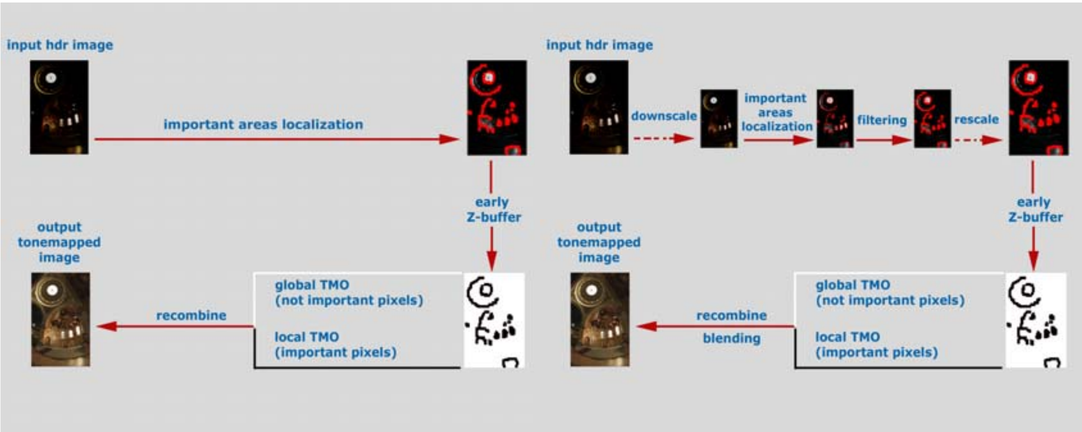

Fig. 1: Scheme representing the general behavior of our Framework (left), and the detailed version (right).

Framework

Here we describe the basic framework used to exploit the idea of the *Selective Tone Mapper*, fig. 1.

The framework comprises the following steps:

Important Areas Localization

An edge filtering method is used to detect the edges in the frame. If a non edge is generated it can be further eliminated in order to maximize the use of the early Z buffer test.
In fig. 1 is shown that this step can be applied either on a scaled size frame (dash line) or on the entire frame. In case it is applied on a scaled size frame a rescale of the frame is requested before to apply the marking step (next step).

Also other techniques that make use of more complex visual attention aspects of the HVS as saliency map, task map or combination of the two maps can be used in this step.
In the applications where the user interacts with an object in the scene he focuses on that object, so important areas are the pixels that are covered by this object. It will be easy to mark these important areas and we only have to render on the texture that holds the object of interest.

Marking Step

This step is required to mark the perceptually important areas of the frame and drive properly the application of the TMO as described in the next step.
For each frame the Z-buffer is prepared with the mask, computed in the previous step.
Early Z rejection is used to tell which pixel is going to be computed, which is similar to using a conditional test for each pixel.

Tone Mapping Application

This step directly applies the two components of the original TMO such as global and local.
Making use of the information obtained from the previous step the local component is applied only on the important areas, instead the global operator is applied on the remaining areas of the frame.
There is a significant saving in computation time as the local component of the operator is applied only on the important locations (pixels), while the global component of the operator is applied on the remaining pixels. The outputs of the application of these two components are afterwards re-combined to obtain the final output image (see fig. 1).

Experimental Results

In this Section we describe the experimental results done for testing the GPU implementation of our Framework. We demonstrate our GPU implementation in a real-time setting; integrating it in a rendering system that receives HDR frames as input. The experiments were conducted on a PC Pentium 4, 3.6 GHz and 2.0 GB of RAM, with graphics card nVidia GForce 6600. We used as pre-existing local TMO a GPU version of the model presented by Ashikhmin [Ash02] [Ro+07].

Quality Comparisons

We first tested the time performances of the GPU implementation in a real-time setting, varing the frame size with and without downscaling the input frame. In Tables 1 and 2 we show the results of this test (in fps).

HDR Frame	Glo. TM	Loc. TM	Sel. TM
256 x 256	300	49	117
512 x 512	91	13	41
768 x 768	39	6	17
1024 x 1024	26	4	11

Table 1 Results (in fps) of the GPU implementation varying the resolution of the HDR input frame. In this table we compare the fps for the global GPU TMO, local GPU TMO and the Selective GPU TMO. In this case the edge map is applied on the full resolution of the input frame.

In Table 2, it is possible to observe how the framework is able to accelerate the GPU implementation of the original local TMO achieving better computation performances.

HDR Frame	1/2 size	1/4 size	1/8 size	1/16 size
256 x 256	130	171	208	212
512 x 512	44	55	68	75
768 x 768	19	24	28	30
1024 x 1024	12	15	19	21

Table 2 Results (in fps) of the GPU implementation of our framework, applying the edge map on the downscaled input frame.

In Table 2, we want to show how the compuation performances of the framework can be improved just applying the edge map step on the down scaled input frame without changing the quality performances as will be shown in the next paragrapg where a qaulity comparison is presented (see Fig. 6 and Table 7).

The computation performances achieved are comparable with the one obtained by the global TM operator see Table 1. The results of the Tables 1 and 2 are also reported in Fig. 2.

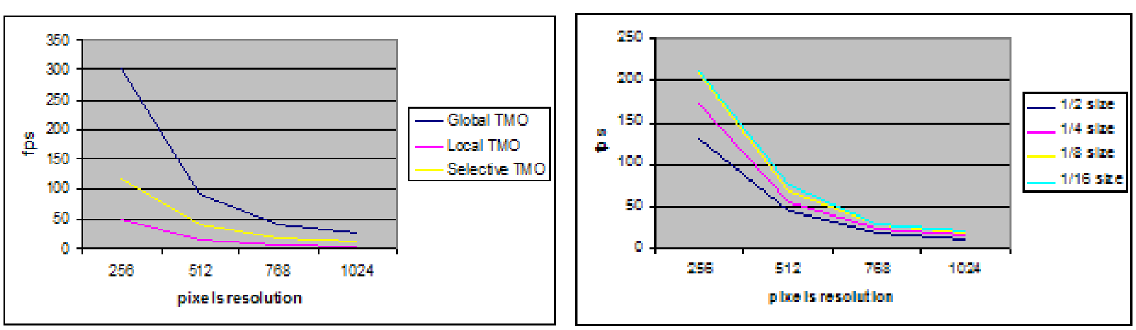

Figure 2 Graphical representation of the Tables 1 and 2.

In order to validate our framework on the latest grapichs technology we perform other tests on the NVIDIA GeForce 8800 GTX.
The results of this evaluation are reported in the Table 3 for the time spent only for the execution of the TM step and Table 4 for the time spent for the rendering plus the TM steps.

HDR Frame	Glo. TM	Loc. TM	Sel. TM
768 x 768	274	68	92
1024 x 1024	237	50	81
2048 x 2048	44	8	14

Table 3 Results (in fps) of the GPU implementation varying the resolution of the HDR input frame. In this table we compare the fps for the global GPU TMO, local GPU TMO and the Selective GPU TMO. In this case the edge map is applied on 1/4 of the resolution of the input frame. The time takes only into account the time spent for the application of the TM operator. The test is done on the new NVIDIA GeForce 8800 GTX.

The results shown in Table 3 confirm that our framework is able to speed-up the computation performances of the original GPU local TMO even on the latest graphics card. In Table 4 we have shown that in case also the rendering step is taken into account the perfromances gained by our framework are close to global TM perfromances.

HDR Frame	Glo. TM	Loc. TM	Sel. TM
768 x 768	74	41	50
1024 x 1024	72	32	46
2048 x 2048	32	7	10

Table 4 Results (in fps) of the GPU implementation varying the resolution of the HDR input frame. In this table we compare the fps for the global GPU TMO, local GPU TMO and the Selective GPU TMO. In this case the edge map is applied on 1/4 of the resolution of the input frame. The time takes into account the time spent for the rendering + the application of the TM operator. The test is done on the new NVIDIA GeForce 8800 GTX.

The next step is to validate the quality reproduction of our framework. To perform this test we have used two well known quality metrics: *Visual Difference Predictor* (*VDP*) [VDP] and *Structural Similarity Index* (*SSIM*) [Wa+04].

The *VDP* gives as output a percentage of pixels that are perceived different from the *Human Visual System* (*HVS*).
The *HVS* has the property to higly adapt to extract structural information from the viewing field. Based on this concept, a measure of structural information changes provide a good approximation to perceive image distortion [Wa+04]. The *SSIM* index provide a measure of the structural information changes in the image. In other words, it indicates the degree of similarities between two images: *1* perfect similarity and *0* completely different.

The use of the Saliency map has been extensively used to significantly reduce the computation of high-fidelity global illumination calculations without deprecating the final quality of the rendered images. We conducted a simple experiment to demonstrate how an edge map is capable to give in our application enough information to drive our Selective Tone Mapper without reducing the quality of the tone mapped images.
The experiment has been performed on different *HDR* images downloadable from the following website [HDR], that span from indor, outdor and daylight. Their characteristics are presented in Table 5.

HDR Image	Resolution	Dynamic Range (log 10)
Memorial	512 x 768	4.8
Belgium	1025 x 768	4.1
Nave	720 x 480	6.0
FogMap	751 x 1330	3.6

Table 5 *HDR* Images used in the experiments. First column is reported the image resolution and in the second column the dynamic range in logarithmic scale.

HDR Image	SSIM	VDP (%)
Memorial	0.9922	2.23
Belgium	0.9793	0.411
Nave	0.9967	5.23
FogMap	0.9972	0.05

Table 6 Results of the quality comparison using the *VDP* and *SSIM* metrics respectively. In the *VDP* metric is resported the percentage of pixels that are perceived different from the *HVS*.

From the results showen in Table 6, the images obtained using the two different maps (Saliency and Edge map) as input for driving the Salective Tone Mapper are not perceived as different from the *HVS*. In fact the *SSIM* index tell that the two images are practically similar and the *VDP* detected a very small percentage of pixels that can be percieved as different form the *HVS*.
From Fig. 3 to Fig. 4 are showen the images obtained using the Edge map (left), the image obtained with the Saliency map (center) and the *VDP* map that is showing the pixels (red dots) perceived as different by the *HVS* (right).

Fig 3 Quality comparison between the images obtained with the use of two different methods for the localization of the important areas. (left) Edge map, (center) Saliency map and (right) VDP map. The VDP map is showing the pixels perceived as different by the HVS (red dots). FogMap image courtesy of Jack Tumblin.

Figure 4 Quality comparison between the images obtained with the use of two different methods for the localization of the important areas. (left) Edge map, (center) Saliency map and (right) VDP map. The VDP map is showing the pixels perceived as different by the HVS (red dots). Belgium image courtesy of Dani Linchinski.

The last experiment performed, is the quality comparison of the outputs obtained using as input the Edge map applied on the original and on the down-scaled size of the input image (see Table 7).

HDR Image	SSIM	VDP (%)
Memorial vs. 1/2	0.9982	0.63
Memorial vs. 1/4	0.9980	0.75
Memorial vs. 1/8	0.9974	0.96
Memorial vs. 1/16	0.9969	1.04

Table 7 Results of the quality comparison using the *VDP* and *SSIM* metrics respectively. In the *VDP* metric is resported the percentage of pixels that are perceived different from the *HVS*.

Conclusions

We introduce the concept of *Selective Tone Mapper*, that is a framework able to accelerate a GPU implementation of pre-existing local TMOs making use of some abilities of the *HVS*. The proposed framework has been included into a real time setting allowing an easy integration into the rendering pipeline. Our method does not require any modification of the rendering pipeline, and proposes an efficient solution that overcomes many of the limitations of current graphics hardware.

Our method shows a speed-up over pre-existing local TMOs, achieving computation performances close to those gained with the corresponding global component of the local operators. We have also shown how making use of an Edge map, in the localization of the important areas, instead of a Saliency map does not change the quality perceived of the output images.

An important feature of the framework is that the GPU implementation does not need to be modified for an application with a different tone mapping operator. The modularity of the proposed framework enables us to concentrate on improving each of its parts independently.

Bibliography

[Ar+03] A. Artusi, J. Bittner, M. Wimmer, A. Wilkie
"Delivering Interactivity to Complex Tone Mapping Operators". Proceedings Eurographics Symposium on Rendering EGSR (2003).

[Ash02] M. Ashikhmin
"A Tone Mapping Algorithm for High Contrast Images". Proceedings Eurographics Workshop on Rendering EGRW (2002).

[Ca+03] K. Cater, A. Chalmers, G. Ward
"Detail to attention: Exploiting Visual Task for Selective Rendering". Proceedings Eurographics Symposium on Rendering EGSR (2003).

[Ca07] M. Cadik
"Perception motivated hybrid approach to tone mapping."Proceedings WSCG 2007.

[Dal93] Daly S.
"The visible differences predictor: an algorithm for the assessment of image fidelity." pp.~189--206. Digital Image and Human Vision.

[devlin2002] Devlin K., Chalmers A., Wilkie A., Purgathofer W.

“Star: Tone reproduction and physically based spectral rendering.” pp.~101--123. Eurographics 2002.

[GWW*03] Goodnight N., Wang R., Woolley C., Humphreys G.
“Interactive time-dependent tone mapping using programmable graphics hardware.” EGSR03, Eurographics Symposium on rendering 2003} (2003),
pp.~26--37.

[IKN98] Itti L., Koch C., Niebur E.
„A model of saliency based visual attention for rapid scene analysis.” Pattern Analysis and Machine Intelligence 20 (1998), pp. 1254--1259.

[Ja90] W. James
“Principle of Psychology”. New York,:Holt.

[kra+05] Grzegorz~Krawczyk Karol~Myszkowski H.-P.~S.
“Perceptual effects in real-time tone mapping with the support of graphics hardware.” Spring Conference in Computer Graphics (2005),
pp.~195--202.

[long06] Longhurst P., Debattista K., Chalmers A.
“Gpu based saliency map for high-fidelity selective rendering.” Afrigraph 2006.

[MR98] Mack A., Rock I.
“Inattentional blindness.” Proceedings of Symposium on Interactive 3D Graphics, Massachusetts Institute of Technology Press (1998).

[My98] Myszkowski K.
“The visible difference predictor: Application to global illumination problems.” Proceeding of Eurographics Workshop on Rendering, G. Drettakis and N. Max Eds. (1998), pp.~223--236.

[RE02] R. Rensink
“Visual attention” In Encyclopedia of Cognitive Science, London: Nature Publishing Group, 2002.

[Ro+07] Roch B., Artusi A., Michael D., Chrysanthou Y., Chalmers A.
“Interactive local tone mapping operator with the Support of Graphics Hardware.” Spring Conference in Computer Graphics, Proceedings SCCG 2007.

[San] P.V.Sander, J.R.Isidoro, Mitchell J.
“Computation culling with explicit early-z and dynamic flow control.”

[Su+04] V. Sundstedt, A. Chalmers, K. Cater and K. Debattista

"Top-Dowm Visual Attention for Efficient Rendering of Task Related Scenes". In: *VMV 2004 - Vision, Modelling and Visualization*, Stanford, November 2004.

[SC99] Simons D., Chabris C.
"Gorillas in our midst: Sustained inattentional blindness for dynamic events."Perception 28 (1999), 1059--1074.

[Yee+01 Yee H., Pattanaik S., Greenberg D.
"Spatiotemporal sensitivity and visual attention for efficient rendering of dynamic environments." ACM Transaction on Computer Graphics 20}, 1 (2001), 39--65.