TAMPERE UNIVERSITY OF TECHNOLOGY

Sumeet Sen

Implementation of Depth Map Filtering on GPU

Master of Science Thesis

Examiners: Prof. Atanas Gotchev

Examiner and topic approved in
the Department of Signal Processing
meeting on xx.xx.xxxx

ABSTRACT

TAMPERE UNIVERSITY OF TECHNOLOGY
Master's Degree Programme in Information Technology
Sumeet Sen : Implementation of Depth Map Filtering on GPU
Master of Science Thesis, xx pages, x Appendix pages
May 2010
Major: Signal processing
Examiner: Prof. Atanas Gotchev
Keywords: General Purpose Graphics Processing Unit (GPGPU), OpenCL, Bilateral Filter, Hypothesis Filtering

First paragraph: Why use depth-map? All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy...All work and no play makes Jack a dull boy, all work and no play makes Jack a dull boy, all work and no play makes Jack a dull boy.

Second paragraph: Why GPU and OpenCL? All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy...All work and no play makes Jack a dull boy, all work and no play makes Jack a dull boy, all work and no play makes Jack a dull boy.

Third paragraph: How did we improve the situation ? How did we measure the improvement. All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy.

PREFACE

The preface presents the general information concerning the thesis process. It describes the supervisor's role and provides notes on the outside help the author received while doing the work (instructions, guidance, etc.). With regard to team work, the author's role in comparison to the other writers is described. Acknowledgements to the persons and communities influencing the work process are generally presented in the preface. It is not appropriate to criticise anyone in the preface. The maximum length of the preface is one page. The date after which no revisions have been made, the author's signature, name in block letters and, if so desired, the address and other contact information are presented at the end of the preface.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABBREVIATIONS AND NOTATION

GPGPU	General Purpose Graphics Processing Unit
GPU	Graphics Processing Unit
CPU	Central Processing Unit
OpenCL	Open Computer Language
2D	Two Dimension
3D	Three Dimension
CV	Computer Vision
HVS	Human Visual System

1. INTRODUCTION

1.1 Problem definition

The Graphics Processing Unit, better known as GPUs came into existence to satisfy the ever increasing demand for computational power. It has proved to be especially effective in the field of computer graphics where the work load is parallel in nature i.e. same operations are performed on large set of data. One such field of research is the 'stereoscopy' where we create or enhance the illusion of depth.

This illusion is a result of viewing two different views of the same image by our eyes; which are reconstructed in the brain to give us an immersive experience or the 3D experience. The three most important components for 3D are the left-view channel, the right-view channel and the depth-map. The left and right views are color images that are captured through two different lenses separated by a small distance. Depth-map is a gray-scale image which represents the distance of the corresponding color pixel from the view-point.

There are different ways for obtaining the depth-maps for a scene. The most commonly used are the 'depth-from-stereo', 'depth-from-multi-view' or depth estimated using sensors. Even though these technologies are promising, there are few common limitations associated with them. The depth computed is sometimes inaccurate, noisy, of low resolution or even inconsistent over a video sequence.

In spite of such drawbacks, reconstructing the other views using the depth image has significant advantages when we take into account the transmission of 3D content. Since depth-map is a gray scale image, it can be compressed significantly compared to a color image (left or right view).

Now, compression adds its own characteristic artifacts (like blockiness) to the depth image. Thus, it can be seen that on the whole, the depth image suffers from inaccuracies, noise and compression artifacts. In case the view is rendered from such a depth-map, it will be crippled with severe distortions.

1.2 Objective

In this work, I tried to enhance the quality of such depth-maps at real time using GPUs. Bilateral filter was used for the iterative refinement process, using the information from one of the high quality texture (color) image i.e. the left or right view.

In the proposed filtering method, we estimate a cost volume and then iteratively refine every slice to obtain the final depth map. The cost volume is built based on the current depth map. Bilateral filtering is then performed throughout each slice of the cost volume to produce a new cost volume. The refined depth map is generated based on the cost volume by selecting the depth hypothesis with the minimal cost.

In order to make the implementation more efficient, some modifications were proposed in the later works. It was observed that there is no need to form cost volume in order to obtain the depth estimate for every pixel and all iterations. Instead, the cost function can be formed only for required neighborhood before filtering it. Furthermore, the not all depth hypothesis were applicable for a given pixel. This computation cost can be reduced by taking into consideration only depth within certain range. Now, in case of compressed depth maps with blocky artifacts the depth range appears to be sparse due to the quantization effect. Such depth maps were scaled to further reduce the number of hypothesizes.

The MEX implementation of this algorithm showed satisfactory results for both quantitative and qualitative experiments. Thus it was quite relevant for the mobile 3D TV kind of setup. My aim was to make the filtering work in real-time and hence we decided to port it on to a GPGPU (General Purpose Graphics Processing Unit).

1.3 Research questions

- It is quite obvious that all algorithms can not be parallelized. We intend to find out whether filtering algorithms using Bilateral filter can attain desires speed-ups when implemented on GPU.

- The ease of implementation of filtering algorithms on GPU using open specifications like the OpenCL.

- The possible boost in computation efficiency that can be attained for OpenCL implementation over more traditional C code running on CPU.

1.4 Limitations

The OpenCL implementation provided in this thesis in not the most efficient one. The memory usage can be further optimized to extract better performance off the GPU. I did put it considerable efforts in this direction. However, it made the code unstable and inconsistent.

1.5 Contribution

- I have tried to put forward the recent techniques to utilize the computing power of our computers which might contain single or multiple CPUs and GPUs. For this purpose I have used device agonistic and open source platform called OpenCL.

- It will provide the reader a basic tutorial to start working with OpenCL.

- It also contains some of the tools which one might find useful while using OpenCL

1.6 Structure of this thesis

The technical, theoretical and other background information which is necessary in order to understand the solutions and methods described later in the thesis is presented in this chapter. The contents of the chapter vary according to the field and study type.

2. A BRIEF INTRODUCTION TO STEREOSCOPY

The theoretical section serves to present a thorough literature review and build the theoretical framework for the entire study. Thorough concept definition is essential for the reader to understand what the author intends. It is important that the theoretical section forms a coherent whole with the entire thesis. Secondary issues concerning the actual thesis are not presented in the theoretical section. The reader can be assumed to be a technology professional without any specialized knowledge of the topic in question. Hence, it is not necessary to present basic issues in the Masters thesis. In addition to the theory, the background information in many cases includes, for instance, a company's old products, the system to be replaced, standards to be followed and a company's internal practices, and other parts of the project if the thesis is part of a larger project. These are presented in as great detail as is necessary to expound the solutions and provide a general view. More detailed outlining and headlining of the chapter (including the main heading) is to be planned and written according to the requirements set by the thesis. The Theoretical background or Starting points chapter can also be divided into several sub chapters.

We humans are gifted with a pair of eyes, which simultaneously captures rays of light and sends the information to our brain. The brain after some complex processing fuses them to create a three dimensional (3D) map of the visible space. This 3D map is unique; it contains the approximate relative distance of all the points in the scene with respect to the eye.

Thus, the photograph captured by a film or digital camera has one fundamental difference with the ones formed in the brain. The former is represented two dimensions (2D) i.e. height and width. It lacks the depth information, which in the field of computer vision (CV) is regarded as the third dimension.

Since, 3D viewing requires at least two light sensors, it is also regarded as stereoscopy. However, study [12] suggests, along with stereo vision the brain might use other cues to determine the relative distances and depth in a scene. These cues are gathered by only one eye and hence known as monocular (or extrastereoscopic) cues. They play a significant role in perception of depth. To an extent, the significance depends on the distance between the observer and the scene. Some of the important ones are listed below.

- Accommodation of the eye: Human eye can change its optical power in order to focus on the object of interest. This is achieved by changing the form of the elastic lens. Even though this is mostly a reflex action but can be controlled to some extent manually.
- Relative size: Objects which creates a bigger image on the retina is perceived to be closer.
- Relative height: Objects placed higher in the scene generally tend to be perceived as further away.
- Overlapping of objects: If one object is occluded by another, it is obvious that the occluding object is closer. This is also known as interposition.
- Linear perspective: The image of two parallel lines or edges, formed on the retina, seem to converge in such a way that they will eventually meet at infinity.
- Blur and desaturation: When the eyes focus on an object it forms a sharp and a color rich image on the retina. The part of the scene that lies outside the focus plane appears to be blur. Desaturation i.e. converting to gray-scale, a part of a scene also gives the effect of out-of-focus region.
- Haze, De-saturation, and a shift to bluishness: In the real world the light traveling from distant objects has to pass through dust or water vapor present in the air giving rise to such effects.

- Other cues: Light, cast shadow and textured pattern are some of the cues commonly used to replicate 3D scene.

2.1 Stereoscopic 3D

In this thesis, we will deal with stereoscopic images which are actually had the third dimension i.e. depth or disparity map. Such images are captured with two cameras which are synchronized and separated by a certain distance. Thus we obtain two images, left-view and right-view.

A disparity map is computed from these two images. Each point may be considered as a two dimensional vector, which maps the corresponding points in the left and right view of a stereo pair. This is more commonly known as stereo matching or stereo correspondence and is discussed by D. Scharstein and R. Szeliski in [13].

Using the method of triangulation, a depth map is computed which indicates the relative distance of a point in the scene from the camera plane. Considering the disparity as d, base offset i.e. separation between the two cameras as b and focal length to be f, the depth D can be calculated as

$$D = \frac{b \cdot f}{d} \tag{2.1}$$

The intensity values of depth-map in the *figure 2.1* is directly proportional to the distance of the region from the camera. Where as, the disparity

2.2 Depth-Image Based Rendering

All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy...All work and no play makes Jack a dull boy, all work and no play makes Jack a dull boy, all work and no play makes Jack a dull boy.

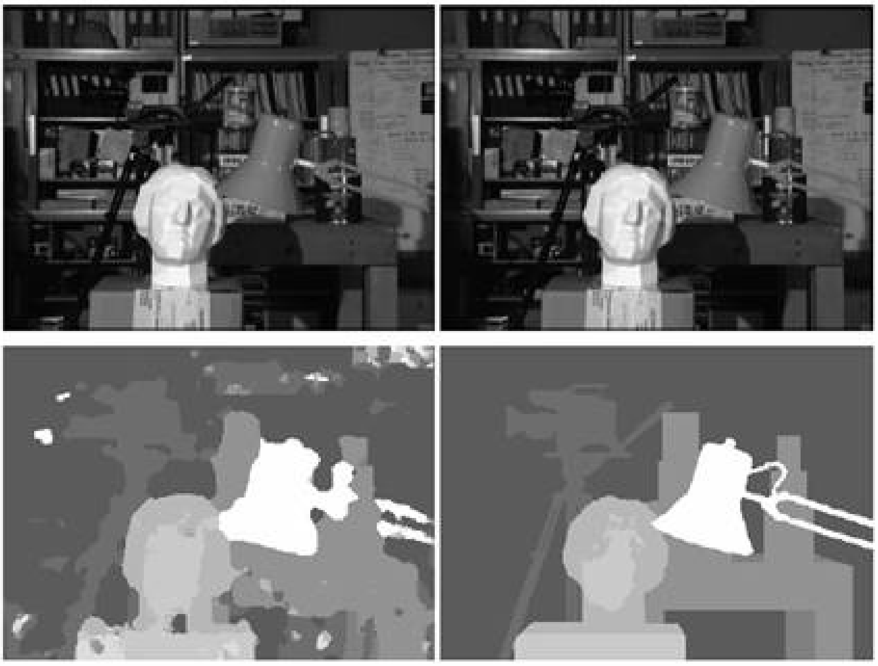

Figure 2.1: *Top left and top right forms the two views of stereo pair. Lower left shows the disparity-map and lower right shows the depth-map [14] .*

2.3 Encoding of depth map

All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy...All work and no play makes Jack a dull boy, all work and no play makes Jack a dull boy, all work and no play makes Jack a dull boy.

2.4 Issues related to Depth-maps

All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy...All work and no play makes Jack a dull boy, all work and no play makes Jack a dull boy, all work and no play makes Jack a dull boy.

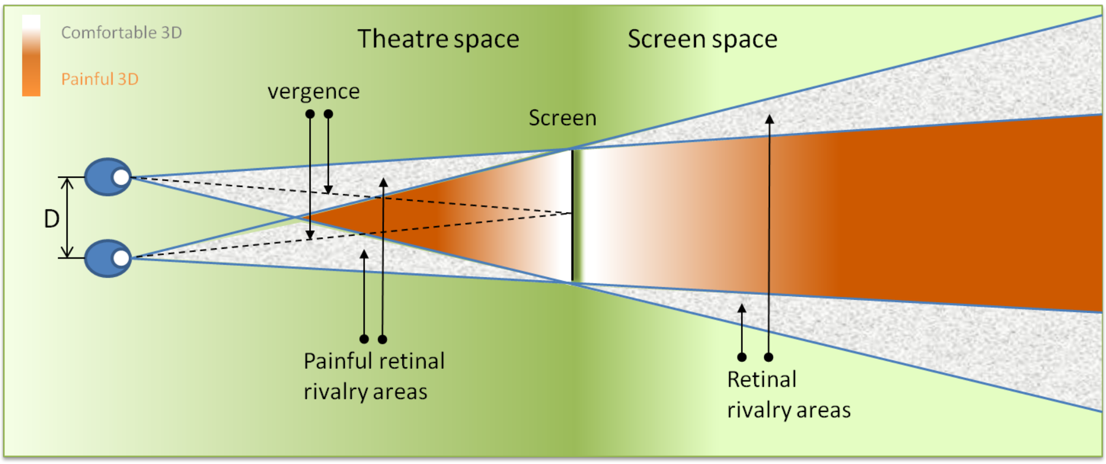

Figure 2.2: *something... [11]* .

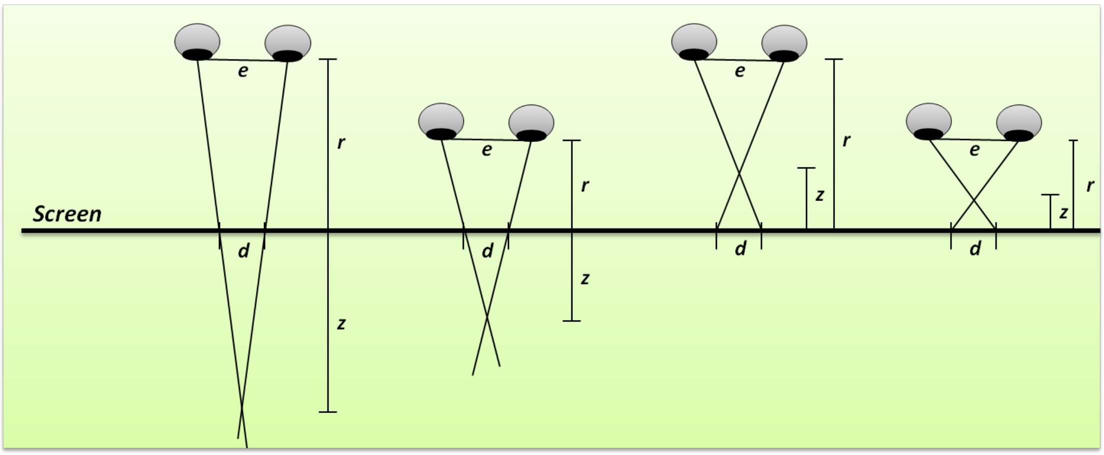

Figure 2.3: *depth Vs viewing distance... [11]* .

2.5 Effects of improper 3D

Improper 3D production will stress the viewer i.e. bad viewer experience

2.6 Filtering of depth map

All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy...All work and no play makes Jack a dull boy, all work and no play makes Jack a dull boy, all work and

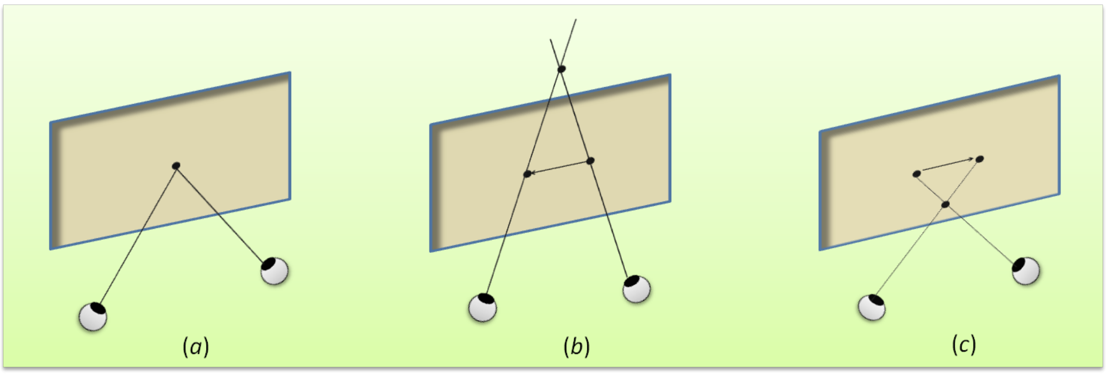

Figure 2.4: *depth parallax... [10]* .

no play makes Jack a dull boy.

2.6.1 Spatial-Depth super resolution for Range Images

As discussed in [?], using one or two registered color images as reference, we can iteratively refine the low-resolution depth map, in terms of both its spatial resolution and depth precision. This whole process is illustrated in the *figure 2.6*.

The iterative refinement process has three major steps. First, a cost volume C_i is generated from the current depth-map $D_{(i)}$. Second, a Bilateral Filtering is performed for each slice of the cost volume to generate a new cost volume $C^{CW}_{(i)}$. Finally the depth hypothesis with minimal cost is selected and a sup-pixel estimation is applied to obtain the final depth-map $D_{(i+1)}$.

Bilateral Filter forms one of the integral part in the iterative refinement process. It has been previously used with great success for stereo algorithms in [?] and [?]. In the next chapter we will elaborately discuss on this filtering scheme and also show how it's performance can be enhanced using a GPU device.

2.7 Transmission of stereo content

All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy...All work and no play

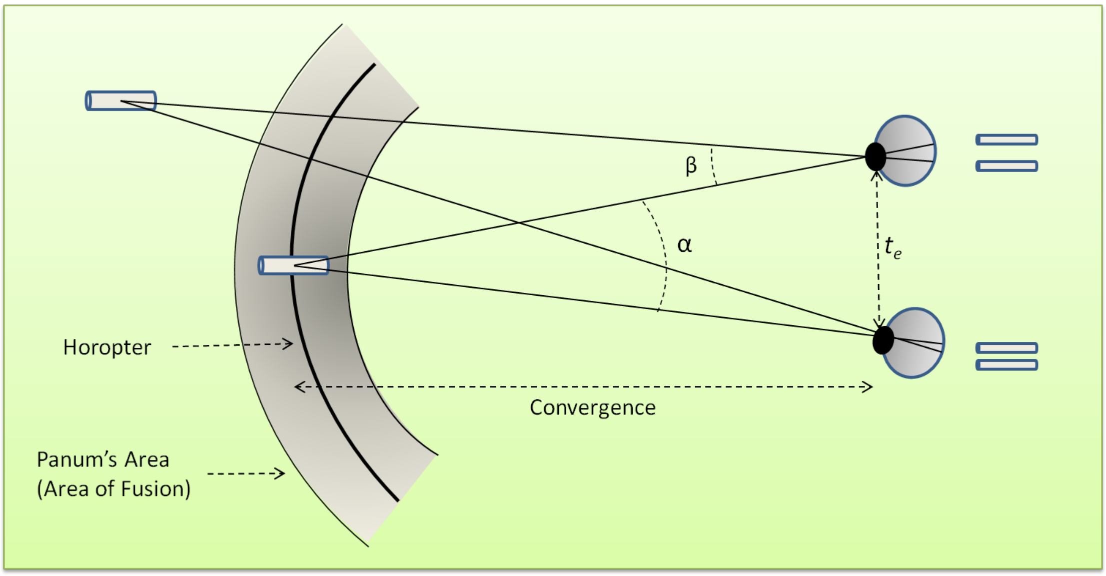

Figure 2.5: *retinal disparity... [10]* .

makes Jack a dull boy, all work and no play makes Jack a dull boy, all work and no play makes Jack a dull boy.

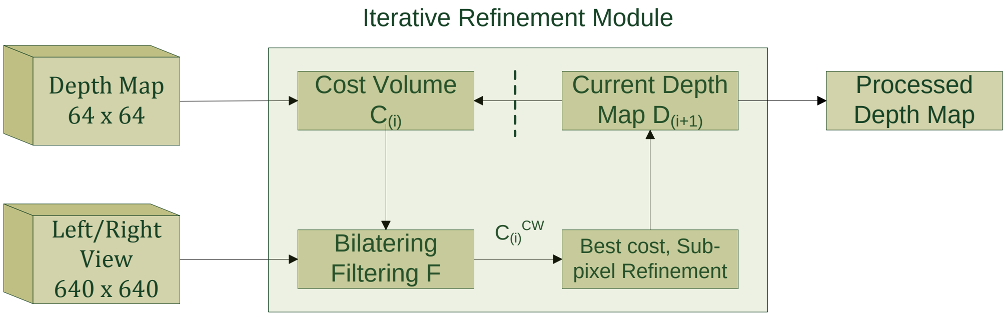

Figure 2.6: *Framework for spatial-depth super resolution for range images. The range image serves as the initial depth-map hypothesis. Then follows an iterative refinement process. A cost volume is built according to the current depth map hypothesis. A bilateral filter is then applied to the cost volume to avoid the smoothening effect near the depth discontinuities. A winner-takes-all and sub-pixel estimation procedure is used to produce a new depth map hypothesis, which is fed back into the process as described in [**?**].*

3. OPENCL: THE OPEN SOURCE SPECIFICATION FOR GPU PROGRAMMING

OpenCL stands for Open Computing Language initially proposed by Apple in 2008 and is currently maintained by Khronos Group. The other big companies involved in developing the OpenCL specification include NVIDIA, AMD, Intel etc.

OpenCL is a specification i.e. programmers has to write their own implementations which are compliant to the specification. Thus OpenCL is a programming interface which offers a framework to build applications on top of it. This framework allows the user to take the advantage of all the system resources (CPU, GPU, DSP chip etc).

OpenCL is designed to support general purpose parallel computation. It can be used for variety of tasks which involves heavy computation and calculation like in computer graphics, digital signal processing, scientific calculations, analysis of financial data and many more.

3.1 The OpenCL Architecture

All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy...

3.1.1 Platform Model

The platform model consists of a host connected to one or more OpenCL devices. The devices in turn contains compute units (CUs) which is a collection of processing units (PEs) which are responsible for all the computations happening in the device. This is shown in the *figure 3.1.* The developer submits the commands to be executed on the device through the host which is responsible for setting up the device. The device splits up the instructions and data on its processing units to be executed as SIMD or SPMD.

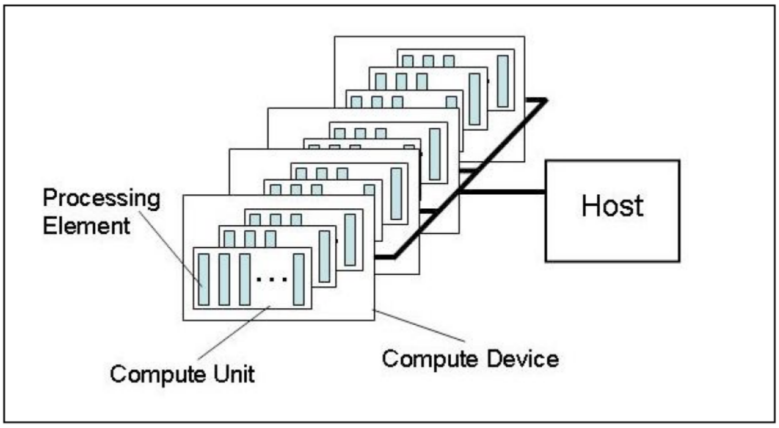

Figure 3.1: *Platform Model ... one host, plus one or more compute devices, each with one or more compute units, each with one or more processing elements [?, p. 23].*

3.1.2 Execution model

Programming in OpenCL has two basic parts: the **host code** that executes on the CPU or the host and the **kernel codes** that execute on one or more OpenCL device. The way in which the kernel executes on the work units of the device defines the execution model in OpenCL.

The basic element of work unit is called **work-item** which are grouped to form **work-groups**. All the work-groups are put together to form the **NDRange** i.e. n-Dimensional Range. The total number of the executable units on the GPU is called the **global size** which is the size of the NDRange. On the other hand, the size of a work-group is called the **local size**. It is to be noted that on a GPU, the global size needs to divide evenly into local size.

One of the powerful feature of OpenCL, is the natural ability to index the global (i.e. NDRange) and local work space in one, two and three dimensions, using the inbuilt functionality provided by the specification. This concept was well explained in [?, p. 24-25] which is also mentioned below.

For example, consider the 2-dimensional index space in *figure 3.2*. We input the index space for the work-items (G_x, G_y), the size of each work-group (S_x, S_y) and the global ID offset (F_x, F_y). The global indices define an G_x by G_y index space where the total number of work-items is the product of G_x and G_y. The local indices define a S_x by S_y index space where the number of work-items in a single work-group is the product of S_x and S_y. Given the size of each work-group and the total number of work-items we can compute the number of work-groups. A 2-dimensional index space is used to uniquely identify a work-group. Each work-item is identified by its global ID (g_x, g_y) or by the combination of the work-group ID (w_x, w_y), the size of each work-group (S_x, S_y) and the local ID (s_x, s_y) inside the work-group such that

$$(G_x, G_y) = (w_x \times S_x + s_x + F_x, \quad w_y \times S_y + s_y + F_y) \tag{3.1}$$

The number of work-groups can be computed as:

$$(W_x, W_y) = (G_x/S_x, \quad G_y/S_y) \tag{3.2}$$

Given a global ID and the work-group size, the work-group ID for a work-item is computed as:

$$(w_x, w_y) = ((g_x - s_x - F_x)/S_x), \quad (g_y - s_y - F_y)/S_y) \tag{3.3}$$

3.1.3 Memory Model

The memory model can be classified into four address spaces which are described in *figure 3.3*. More details on the allocation and usage of the memory available in the device can be found from the *table 3.1*.

- **__global**: It refers to the **global memory** which allows read and write access to all work-items in the work-group.

- **__constant**: It refers to the **constant memory** which is a part of the global memory. However it remains constant during the execution of the kernel. The initialization and the allocation of the memory objects for the constant memory is done by the host.

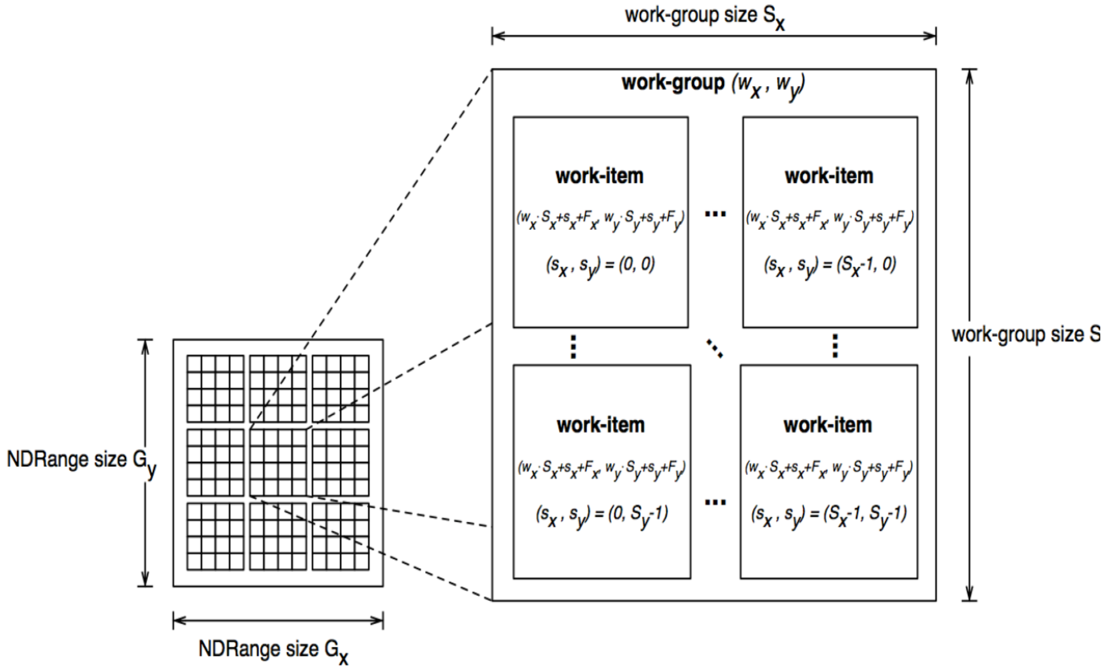

Figure 3.2: *An example of NDRange index space showing work-items, their global IDs and their mapping onto the pair of work-groups and local IDs [?, p. 25].*

- **__local**: It refers to the **local memory** of the work-group which can be shared by all the work-items belonging to that work-group. They are much faster compared to the global memory even though their availability is quite limited.

- **__private**: It refers to the **private memory** of the work-item. Private memory of one work-item is unavailable to other work-items. They act mostly like the resistors and hence are extremely fast.

3.1.4 Programming Model

All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy...

	Global	**Constant**	**Local**	**Private**
Host	Dynamic allocation Read/Write access	Dynamic allocation Read/Write access	Dynamic allocation No access	No allocation No access
Kernel	No allocation Read/Write access	Static allocation Read-only access	Static allocation Read/Write access	Static allocation Read/Write access

Table 3.1: *Memory Region - Allocation and Memory Access Capabilities [**?**, p. 28]*

3.1.5 Memory Objects

All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy...

3.1.6 The OpenCL Framework

All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy,All work and no play makes Jack a dull boy...

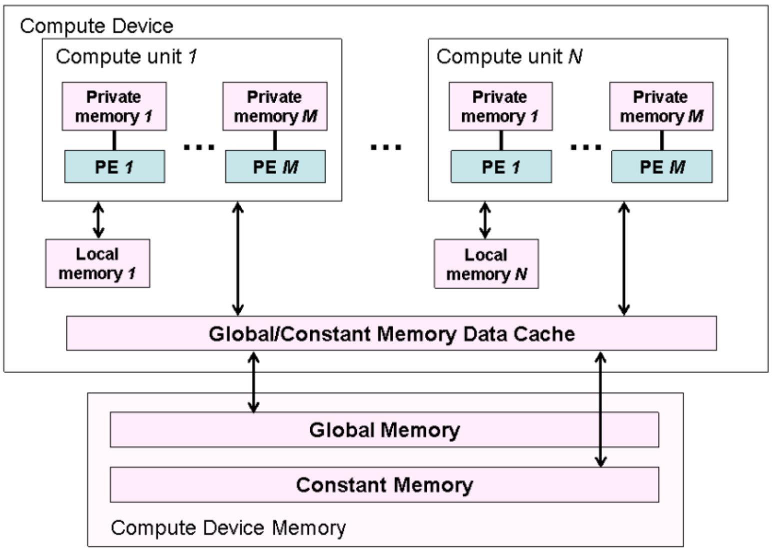

Figure 3.3: *Conceptual OpenCL device architecture with processing elements (PE), compute units and devices [*?*, p. 28].*

3.2 The OpenCL Platform Layer

3.2.1 Querying Platform Info

3.2.2 Querying Devices

3.2.3 Partitioning a Device

3.2.4 Contexts

3.3 The OpenCL Runtime

3.3.1 Command Queues

3.3.2 Buffer Objects

3.3.3 Image Objects

3.3.4 Sampler Objects

3.3.5 Program Objects

3.3.6 Kernel Objects

3.3.7 Executing Kernels

3.3.8 Flush and Finish

3.4 The OpenCL C Programming Language

3.4.1 Supported Data Types

3.4.2 Conversions and Type Casting

3.4.3 Operators

3.4.4 Vector Operations

3.4.5 Address Space Qualifiers

3.4.6 Access Qualifiers

3.4.7 Function Qualifiers

3.4.8 Preprocessor Directive and Macros

3.5 Limitations of OpenCL

Memory Optimizations

NDRange Optimizations

Instruction Optimizations

Control Flow

4. ALGORITHM AND IMPLEMENTATION

The purpose of the research method and material description is to provide the reader with a clear view on the implementation process for the reader to assess its reliability. First, it may be useful to present possible alternative research methods and provide grounds for the reasons behind the method selection in the thesis. The used material should be presented in detail so that there can be no confusion concerning its origin and nature.

The methods should be described in such detail that other researchers in the field could reproduce the study. For instance, the mathematical grounds for new results must be presented in such detail that the reader can follow the result process without having to make long separate calculations. A brief note or reference is sufficient with regard to generally known methods; more unusual methods, particularly those developed by the author, are to be described in more detail. There should always be a clear connection between the theoretical section and the research methods and material. If necessary, the research methods and material can be dealt with in several chapters. The outlining and headlining should be made in accordance with the nature and needs of the thesis.

4.1 Depth map filtering approach

We consider color video sequence in YUV color space $y(x,t) = [y^Y(x,t) \;\; y^U(x,t) \;\; y^V(x,t)]$

4.1.1 Bilateral filter

The purpose of this filter is to smooth the image while preserving the edges [?]. It utilizes the information form all color channels to specify suitable weights for local (non-linear) neighborhood filtering. For gray-scale images, local weights of neighbors are calculated based on both their spatial distance and their photometric similarity, favoring nearer values to distant ones in both spatial domain and intensity range. For color images, bilateral filtering uses color distance to distinguish photometric similarity between pixels, thus reducing phantom colors in the filtered image [?].

$$\hat{Z}(k) = \frac{1}{F} \sum_{m \in \Gamma(k)} e^{-\frac{\|m-k\|^2}{2\sigma_s^2}} e^{-\frac{|Z(m)-z(k)|}{2\sigma_i^2}} \tag{4.1}$$

where σ_s and σ_i controls the weights in spatial and intensity domains.

$$F = \sum_{m \in \Gamma(k)} e^{-\frac{\|m-k\|^2}{2\sigma_s^2}} e^{-\frac{|Z(m)-z(k)|}{2\sigma_i^2}} \tag{4.2}$$

where F is the normalization factor and $\Gamma(k)$ is a square window centered at $Z(k)$.

4.1.2 Hypothesis filter

$$C_{(i)}(x,d) = \min(\delta * L, (d - \hat{Z}_{(i)}(x))^2) \tag{4.3}$$

$$\widehat{C(i)}(x,d) = r(x) \sum_{u \in \Omega_x} W_s(\|x-u\|) W_c(|y(x)-y(u)|) C_{(i)}(u,d) \tag{4.4}$$

$$\hat{Z}_{(i+1)}(x) = arg \min_d (\widehat{C(i)}(x,d)) \tag{4.5}$$

4.2 Implementation in OpenCL

4.2.1 Installing OpenCL

4.2.2 Example: Box Filter implementation in OpenCL

Illustration

Basic steps for OpenCL programming

The OpenCL code

4.2.3 Implementation of Bilateral Filter

4.2.4 Implementation of Hypothesis filter

5. RESULTS AND ANALYSIS

The purpose of the Results and Discussion chapter is to present the results of the study and their significance to the reader. It may be useful to divide the Results and Discussion chapter into several subchapters and provide them with titles which describe the contents in detail. The most important results, their error sources, deviations from the assumed results and reliability of the results are presented in the Results section. The results are to be presented briefly and with precision. Images and tables can be used to aid the understanding of the results. The central message of the images and tables must also be presented in the text. If necessary, supplementary information can be presented in appendices. The results must be understandable without having to read the other sections of the thesis in detail. In the Discussion section, the results are to be compared to the previous study presented in the theoretical section. It is also to be assessed whether the results correspond with the set objectives, as well as does the thesis answer the set research questions. New information or otherwise significant information is to be emphasised in the discussion. In addition, the scientific and practical significance of the results is to be assessed.

6. CONCLUSION

6.1 Conclusion

The Conclusion is one of the most important chapters in the thesis. Many readers often read only the introduction and conclusion. Do not rely on the results to speak for themselves; the authors task is to show the contribution of the thesis to the reader with thorough explanations. Individual results are not repeated here, instead the main results are collected and their meaning is reflected. On the basis of the study, the author presents concrete procedures and recommendations concerning the application of the results for different usage purposes and assesses the limitations regarding their use. The recommendations can be focused directly on the client or the field and society in general. The author should also assess the need for further research or development and the success of the work.

6.2 Future Works

REFERENCES

[1] C. Tomasi, R. Manduchi, "Bilateral Filtering for Gray and Color Images", iccv, pp.839, Sixth International Conference on Computer Vision (ICCV'98), 1998

[2] Sergey Smirnov, Atanas Gotchev, Karen Egiazarian, "METHODS FOR RESTORATION OF COMPRESSED DEPTH MAPS: A COMPARATIVE STUDY", Fourth International Workshop on Video Processing and Quality Metrics for Consumer Electronics (VPQM09), 2009

[3] Sergey Smirnov, Atanas Gotchev, Sumeet Sen, Gerhard Tech, Heribert Brust, "3D Video Processing Algorithms Part I", MOBILE3DTV Project No. 216503

[4] Sergey Smirnov, Atanas Gotchev, Karen Egiazarian, "A Memory-Efficient and Time-Consistent Filtering of Depth Map Sequences", Society of Photo-Optical Instrumentation Engineers (SPIE) Conference Series, FEB 2010

[5] Qingxiong Yang, Ruigang Yang, James Davis, David Nist´er, "Spatial-Depth Super Resolution for Range Images", University of Kentucky University of California, Santa Cruz

[6] Q. Yang, L. Wang, R. Yang, H. Stew´enius and D. Nist´er, "Stereo Matching with Color-Weighted Correlation, Hierarchical Belief Propagation and Occlusion Handling" Center for Visualization and Virtual Environments, Department of Computer Science, University of Kentucky

[7] K-J. Yoon, S. Kweon, "Adaptive Support-Weight Approach for Correspondence Search", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 28, no. 4, APRIL 2006

[8] Aaftab Munshi "The OpenCL Specification", Khronos OpenCL Working Group, Version 1.2, pg 22

[9] K. Dabov, A. Foi, V. Katkovnik, K. Egiazarian, "Image Denoising by Sparse 3-D Transform-Domain Collaborative Filtering", IEEE Transactions on Image Processing, Vol 16, AUG 2007

[10] Frederik Zilly, Josef Kluger, Peter Kauff, "Production Rules for Stereo Acquisition", Proceedings of the IEEE, Vol. 99, No. 4, April 2011

[11] Aljoscha Smolic, Peter Kauff, Sebastian Knorr, Alexander Hornung, Matthias Kunter, Marcus Muller, Manuel Lang, "Three-Dimensional Video Postproduction and Processing", Proceedings of the IEEE, Vol. 99, No. 4, April 2011

[12] J. M. Rolfe and K. J. Staples, Flight Simulation, Cambridge University Press, 1986, page 134

[13] D. Scharstein and R. Szeliski, "A taxonomy and evaluation of dense two-frame stereo correspondence algorithms", International Journal of Computer Vision, April June 2002

[14] M. Dominguez Morales, A. Jimenez Fernandez, R. Paz Vicente, A. Linares Barranco, G. Jimenez Moreno, "Current Advancements in Stereo Vision", book edited by Asim Bhatti, Chapter 1

A. APPENDIX