MULTIMEDIA INFORMATICS
MMI 713– Applied Parallel Programming on GPU
2020-2021 Spring

S Y L L A B U S

Location & Time:　　Online (link will be announced on ODTUClass)- Monday 13:40 - 16:30
Instructor:　　　　　Prof.Dr. Alptekin Temizel, **atemizel@metu.edu.tr**

COURSE OUTLINE

Synchronous sessions are marked in Red. There will be extra synchronous Q&A sessions at some lecture times (Monday 13:40) and these will be announced in due course.
The course content will be uploaded as YouTube videos (+ presentations on ODTUClass).

Week	Date	
1	15 Mar	Introduction to GPUs and GPU Programming
		History of GPU and General Purpose Programming on GPU (GPGPU)
		GPU based applications, programming frameworks
		Advantages and disadvantages of GPGPU
2	22 Mar	Introduction to GPU Architecture and CUDA Basics
		PC and GPU architecture
		Introduction to CUDA, CUDA API and SDK
		CUDA Threads, Code Walkthrough
3	29 Mar	Hands-on Lab
		CUDA environment installation
		Implementation of sample programs
4	5 Apr	Effective Use of Memory and Optimization
		Global Memory, coalescing
		Shared Memory, bank conflicts
		Latency Hiding
		Page-locked transfers, Asynchronous access
5	12 Apr	Effective Use of Memory and Optimization – Lab
		Zero Copy
		Constant Memory, Read-only memory
		Textures
6	19 Apr	Control Flow, parallel reduction, optimization
		Block partitioning, warps,
		Branch divergence overhead
		Parallel Reduction
		Floating point and double operations
		Atomic Operations
7	26 Apr	Project Proposals
		Presentation of project proposals by the students
8	3 May	CUDA Debugging, Profiler/Parallel Nsight
		Local and remote debugging
		PTX and SASS debugging
		Performance analysis and profiling
		Parallel Nsight statistics and examples
	Spring Break 8-16 May	
9	17 May	Analytical Modeling of Parallel Systems and Performance Analysis
		Analytical modelling
		Overheads in parallel programs, Performance Metrics
		Cost of a parallel system
		Amdahl's law
		Scaling characteristics of parallel systems

10	24 May	Thrust Library
		Thrust Basics
		Templates and functors
		Thrust algorithms, Optimization in Thrust
11	31 May	OpenCL and Case Studies
		OpenCL Basics
		Case Study I – Background Subtraction
		Implementation of background subtraction algorithm on GPU
		Memory optimization (coalesced access, asynchronous copying)
		Case Study II – Cross Correlation
		Implementation of cross correlation on GPU
		Memory optimization (shared memory, avoiding bank conflicts)
12	7 June	Multi GPU programming and Advanced Concepts
		Using Multiple GPUs
		Dynamic Parallelism
		HyperQ
		WARP Operations and Shuffling
13	14 June	Deployment
		Deployment Tools, GPU containerization
		PyCUDA, Numba
		TensorRT
		Tensor Cores
		GPU Rest Engine
14	21 June	Embedded Platforms
		GPU Programming on Jetson Platforms

REFERENCE MATERIAL

David B. Kirk and Wen-mei W. Hwu "Programming Massively Parallel Processors: A Hands-on Approach – 2nd Edition", Morgan Kaufman, ISBN-13: 978-0-12-415992-1.

H. Bidgoli, "CUDA by Example: An Introduction to General-Purpose GPU Programming", Addison Wesley, ISBN-10: 0131387685, ISBN-13: 978-0131387683.

S. Cook, "CUDA Programming: A Developer's Guide to Parallel Computing with GPUs", ISBN-13: 978-0124159334

Rob Farber, "CUDA Application Design and Development", ISBN-13: 978-0123884268.

NVIDIA Developer Zone, http://developer.nvidia.com/page/home.html

OpenCL Developer Zone, http://www.khronos.org/opencl/

COURSE CONDUCT

The course will be taught in class and also will involve application. The students are expected to present a project proposal that involves parallel programming and implement on GPU using CUDA or OpenCL. The term project will help the students to apply the knowledge obtained in the class and gain practical experience by implementing it.

GRADING

5% Lab Assignment

30% Assignments (3 Assignments – 10% each)

50% Project (Project has 3 phases)

15% Quiz (3 quizzes – 5% each)

5% Attendance – Participation/Quiz

Total: 105%

You will be told about the **quiz time and the content** one week before the quiz at the lecture hours!

No late submissions will be allowed for the assignments and the project phases.

You will submit your work (assignments and project phases) to **ODTUClass**. Note that the system closes at 23:55 at the due date so do not leave it to the last minute.

There are no set office hours, please do not hesitate to contact me on my e-mail **atemizel@metu.edu.tr** to get an appointment if you need help with anything related with the course. Alternatively you can use ODTUClass to ask any question or to share anything related with the course.

The course assistant is Ayberk Aydın – ayberkydn@gmail.com , please contact him regarding any questions on labs or assignments.