



Modern Parallel Programming with C++ and Assembly Language

X86 SIMD Development Using AVX,
AVX2, and AVX-512

—
Daniel Kusswurm

Apress®

Modern Parallel Programming with C++ and Assembly Language

X86 SIMD Development Using AVX, AVX2,
and AVX-512



Daniel Kusswurm

Apress®

Modern Parallel Programming with C++ and Assembly Language: X86 SIMD Development Using AVX, AVX2, and AVX-512

Daniel Kusswurm
Geneva, IL, USA

ISBN-13 (pbk): 978-1-4842-7917-5
<https://doi.org/10.1007/978-1-4842-7918-2>

ISBN-13 (electronic): 978-1-4842-7918-2

Copyright © 2022 by Daniel Kusswurm

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

Trademarked names, logos, and images may appear in this book. Rather than use a trademark symbol with every occurrence of a trademarked name, logo, or image we use the names, logos, and images only in an editorial fashion and to the benefit of the trademark owner, with no intention of infringement of the trademark.

The use in this publication of trade names, trademarks, service marks, and similar terms, even if they are not identified as such, is not to be taken as an expression of opinion as to whether or not they are subject to proprietary rights.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Managing Director, Apress Media LLC: Welmoed Spahr
Acquisitions Editor: Steve Anglin
Development Editor: James Markham
Coordinating Editor: Mark Powers

Cover designed by eStudioCalamar

Cover image by Viktor Forgacs on Unsplash (www.unsplash.com)

Distributed to the book trade worldwide by Apress Media, LLC, 1 New York Plaza, New York, NY 10004, U.S.A. Phone 1-800-SPRINGER, fax (201) 348-4505, e-mail orders-ny@springer-sbm.com, or visit www.springeronline.com. Apress Media, LLC is a California LLC and the sole member (owner) is Springer Science + Business Media Finance Inc (SSBM Finance Inc). SSBM Finance Inc is a **Delaware** corporation.

For information on translations, please e-mail booktranslations@springernature.com; for reprint, paperback, or audio rights, please e-mail bookpermissions@springernature.com.

Apress titles may be purchased in bulk for academic, corporate, or promotional use. eBook versions and licenses are also available for most titles. For more information, reference our Print and eBook Bulk Sales web page at <http://www.apress.com/bulk-sales>.

Any source code or other supplementary material referenced by the author in this book is available to readers on GitHub (<https://github.com/Apress>). For more detailed information, please visit <http://www.apress.com/source-code>.

Printed on acid-free paper

Table of Contents

About the Author	xi
About the Technical Reviewer	xiii
Acknowledgments	xv
Introduction	xvii
■ Chapter 1: SIMD Fundamentals	1
What Is SIMD?	1
Historical Overview of x86 SIMD	4
SIMD Data Types.....	5
SIMD Arithmetic	7
SIMD Integer Arithmetic.....	7
SIMD Floating-Point Arithmetic	10
SIMD Data Manipulation Operations	12
SIMD Programming	15
Summary.....	16
■ Chapter 2: AVX C++ Programming: Part 1.....	17
Integer Arithmetic.....	17
Integer Addition	17
Integer Subtraction	22
Integer Multiplication.....	24
Integer Bitwise Logical and Shift Operations	30
Bitwise Logical Operations	30
Shift Operations	33

C++ SIMD Intrinsic Function Naming Conventions	35
Image Processing Algorithms.....	37
Pixel Minimum and Maximum	37
Pixel Mean Intensity	45
Summary	50
■ Chapter 3: AVX C++ Programming: Part 2.....	53
Floating-Point Operations.....	53
Floating-Point Arithmetic.....	53
Floating-Point Compares	59
Floating-Point Conversions.....	64
Floating-Point Arrays.....	68
Mean and Standard Deviation	69
Distance Calculations	78
Floating-Point Matrices	88
Column Means.....	89
Summary	95
■ Chapter 4: AVX2 C++ Programming: Part 1.....	97
Integer Arithmetic.....	97
Addition and Subtraction	97
Unpacking and Packing	102
Size Promotions.....	107
Image Processing.....	113
Pixel Clipping.....	114
RGB to Grayscale	119
Thresholding.....	128
Pixel Conversions	137
Summary.....	142

■ Chapter 5: AVX2 C++ Programming: Part 2	145
Floating-Point Arrays.....	145
Least Squares.....	146
Floating-Point Matrices.....	151
Matrix Multiplication.....	152
Matrix (4 × 4) Multiplication.....	161
Matrix (4 × 4) Vector Multiplication.....	169
Matrix Inverse.....	181
Summary.....	188
■ Chapter 6: AVX2 C++ Programming: Part 3	189
Convolution Primer.....	189
Convolution Math: 1D.....	189
Convolution Math: 2D.....	192
1D Convolutions.....	194
2D Convolutions.....	206
Nonseparable Kernel.....	206
Separable Kernel.....	215
Summary.....	222
■ Chapter 7: AVX-512 C++ Programming: Part 1	223
AVX-512 Overview.....	223
Integer Arithmetic.....	224
Basic Arithmetic.....	224
Merge Masking and Zero Masking.....	230
Image Processing.....	237
RGB to Grayscale.....	237
Image Thresholding.....	240
Image Statistics.....	247
Summary.....	255

■ Chapter 8: AVX-512 C++ Programming: Part 2	259
Floating-Point Arithmetic.....	259
Basic Arithmetic	259
Compare Operations	265
Floating-Point Arrays.....	269
Floating-Point Matrices	272
Covariance Matrix.....	272
Matrix Multiplication.....	279
Matrix (4 x 4) Vector Multiplication.....	283
Convolutions.....	289
1D Convolutions.....	289
2D Convolutions.....	294
Summary.....	300
■ Chapter 9: Supplemental C++ SIMD Programming	303
Using CPUID.....	303
Short Vector Math Library	315
Rectangular to Polar Coordinates	316
Body Surface Area	325
Summary.....	331
■ Chapter 10: X86-64 Processor Architecture	333
Data Types.....	333
Fundamental Data Types	334
Numerical Data Types.....	335
SIMD Data Types.....	335
Strings	335
Internal Architecture.....	336
General-Purpose Registers.....	337
Instruction Pointer	338
RFLAGS Register.....	338
Floating-Point and SIMD Registers.....	340

MXCSR Register.....	342
Instruction Operands.....	343
Memory Addressing	344
Condition Codes	346
Summary.....	347
■ Chapter 11: Core Assembly Language Programming: Part 1	349
Integer Arithmetic.....	349
Addition and Subtraction.....	350
Multiplication.....	353
Division.....	357
Calling Convention: Part 1	362
Memory Addressing Modes.....	368
For-Loops	373
Condition Codes	376
Strings.....	381
Summary.....	386
■ Chapter 12: Core Assembly Language Programming: Part 2	389
Scalar Floating-Point Arithmetic	389
Single-Precision Arithmetic.....	389
Double-Precision Arithmetic.....	393
Compares	397
Conversions.....	400
Scalar Floating-Point Arrays.....	409
Calling Convention: Part 2	411
Stack Frames.....	412
Using Nonvolatile General-Purpose Registers.....	416
Using Nonvolatile SIMD Registers	420
Macros for Function Prologues and Epilogues	425
Summary.....	431

■ Chapter 13: AVX Assembly Language Programming: Part 1	433
Integer Arithmetic.....	433
Addition and Subtraction	433
Multiplication.....	437
Bitwise Logical Operations	441
Arithmetic and Logical Shifts	443
Image Processing Algorithms.....	444
Pixel Minimum and Maximum	444
Pixel Mean Intensity	448
Summary.....	453
■ Chapter 14: AVX Assembly Language Programming: Part 2	455
Floating-Point Operations.....	455
Floating-Point Arithmetic.....	455
Floating-Point Compares	461
Floating-Point Arrays.....	465
Mean and Standard Deviation	466
Distance Calculations	470
Floating-Point Matrices	477
Summary.....	481
■ Chapter 15: AVX2 Assembly Language Programming: Part 1	483
Integer Arithmetic.....	483
Basic Operations.....	483
Size Promotions.....	486
Image Processing.....	490
Pixel Clipping.....	491
RGB to Grayscale	495
Pixel Conversions	501
Summary.....	504

■ Chapter 16: AVX2 Assembly Language Programming: Part 2	505
Floating-Point Arrays.....	505
Floating-Point Matrices	510
Matrix Multiplication.....	510
Matrix (4 × 4) Multiplication	514
Matrix (4 × 4) Vector Multiplication	518
Signal Processing.....	525
Summary.....	530
■ Chapter 17: AVX-512 Assembly Language Programming: Part 1	533
Integer Arithmetic.....	533
Basic Operations.....	533
Masked Operations.....	537
Image Processing.....	542
Image Thresholding	542
Image Statistics.....	546
Summary.....	552
■ Chapter 18: AVX-512 Assembly Language Programming: Part 2	553
Floating-Point Arithmetic.....	553
Basic Arithmetic	553
Compare Operations	558
Floating-Point Matrices	561
Covariance Matrix.....	561
Matrix Multiplication.....	568
Matrix (4 x 4) Vector Multiplication.....	573
Signal Processing.....	578
Summary.....	586
■ Chapter 19: SIMD Usage and Optimization Guidelines	587
SIMD Usage Guidelines	587
C++ SIMD Intrinsic Functions or x86 Assembly Language.....	588

SIMD Software Development Guidelines	589
Identify Functions for SIMD Techniques	589
Select Default and Explicit SIMD Instruction Sets	589
Establish Benchmark Timing Objectives.....	590
Code Explicit SIMD Functions	590
Benchmark Code to Measure Performance.....	590
Optimize Explicit SIMD Code	591
Repeat Benchmarking and Optimization Steps	591
Optimization Guidelines and Techniques	591
General Techniques	591
Assembly Language Optimization Techniques.....	592
SIMD Code Complexity vs. Performance	594
Summary	602
■ Appendix A: Source Code and Development Tools	603
Source Code Download and Setup	603
Development Tools	604
Visual Studio and Windows	604
GCC and Linux	616
■ Appendix B: References and Resources	621
C++ SIMD Intrinsic Function Documentation.....	621
X86 Programming References	621
X86 Processor Information.....	622
Software Development Tools.....	622
Algorithm References.....	622
C++ References	623
Utilities, Tools, and Libraries.....	624
Index	625