DESIGN AND ANALYSIS OF ARCHITECTURES FOR STEREO VISION

Von der Fakultät für Elektrotechnik und Informatik
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des akademischen Grades
Doktor-Ingenieur
genehmigte Dissertation von

Dipl.-Ing. Christian Banz

geboren am 31. Januar 1983 in Hannover, Deutschland

2013

1. Referent	Prof. Dr.-Ing. Peter Pirsch
2. Referent	Prof. Dr. Tech. Jarmo Takala
Tag der Promotion	12. September 2013

Berichte aus der Informationstechnik

Christian Banz

Design and Analysis of Architectures for Stereo Vision

Shaker Verlag
Aachen 2013

Bibliographic information published by the Deutsche Nationalbibliothek
The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data are available in the Internet at http://dnb.d-nb.de.

Zugl.: Hannover, Leibniz Univ., Diss., 2013

Printed in Germany.

ISBN 978-3-8440-2401-2
ISSN 1610-9406

Shaker Verlag GmbH • P.O. BOX 101818 • D-52018 Aachen
Phone: 0049/2407/9596-0 • Telefax: 0049/2407/9596-9
Internet: www.shaker.de • e-mail: info@shaker.de

ACKNOWLEDGMENTS

This dissertation could never have been completed successfully without the contribution and support of many people. Firstly, I would like to thank my doctoral advisor Prof. Dr.-Ing. Peter Pirsch for his excellent support and guidance from my first days at the Institute of Microelectronic Systems all the way through to my dissertation. Without his trust, several great opportunities would not have been available to me.

I would also like to thank Prof. Dr. Tech. Jarmo Takala for his generous acceptance of the role of Coexaminer. Special thanks goes also to Prof. Dr.-Ing. Holger Blume for being a splendid sparring partner in many discussions, for providing excellent working conditions at the Institute of Microelectronic Systems, and for agreeing to act as the Examining Committee's Chairman.

I am grateful for the funding of the Hans L. Merkle Stiftung. The remarkable programme of the Trust has given me many interesting insights which will stay with me for life. Many thanks also go to the Trust Programme's mentors at Robert Bosch GmbH. Dr. Wolfgang Niehsen, Jochen Wingbermühle, Arne Zender, and Henning von Zitzewitz have given me great scientific and practical advise and always inspired me with new ideas.

I would like to thank all my former colleges at the Institute of Microelectronic Systems for the collegiate atmosphere. Special thanks also go to Prof. Dr. Guillermo Payá Vayá for our collaboration. Further thanks go to Steffen Blume, Daniel Pefferkorn, Sebastian Hesselbarth, Dr. Javier Martín Langerwerf, Nico Mentzer, and Dr. Carsten Dolar for many insightful discussions.

Finally, I would like to thank my parents, Hildegard and Harald Banz, for their constant support and always being there for me; my brother Dr. Lothar Banz for mathematical and other advice at all times; Sheila Isabell White for being there for me when it matters. My deepest thanks go to Dr. Wiebke Schaper for her patience, encouragement, constant support, and for listening until she knew what an FPGA is.

Christian Banz

Hannover, September 2013

ABSTRACT

The demand for high performance image processing systems, driven by ever-evolving image processing algorithms, has begun to grow immensely in recent years. For the implementation of such systems, the designer faces a vast number of choices and has to confront a set of conflicting constraints comprising performance, power consumption, and flexibility, to name just a few. In order to identify optimal solutions there is a prevailing need to constantly evaluate and quantify existing and emerging architectures using relevant, state-of-the-art image processing algorithms. This issue is especially apparent in advanced driver assistance systems, where for many relevant algorithms efficient implementations for demanding throughput constraints are unavailable on any architecture. One of these algorithms is semi-global matching (SGM) which enables 3D perception using inexpensive, passive stereo cameras.

This work presents the design, optimization and analysis of high efficiency implementations in the hardware and software domain of the semi-global matching algorithm. The target is a heterogeneous set of architectures comprising a wide spectrum of current architectural concepts for image processing: two ASPs (one RISC processor and one VLIW processor), a GPU, a multi-core GPP, and a dedicated architecture mapped onto FPGA and standard cell ASIC. A design space exploration (DSE) framework is introduced to assess architectures quantitatively in terms of silicon area, energy consumption, and throughput performance. It contains a specifically designed FPGA-based SoC framework for image processing tasks enabling emulation-based analysis and rapid prototyping. Using the DSE framework, detailed analyses of all respective implementation options for each architecture (intra-architecture analysis) and across architectures (inter-architecture analysis) are performed. The SoC framework and FPGA results are further applied to compose a fully functional real-time stereo vision system.

The evaluation shows that there is an intra-architecture optimization potential in terms of area efficiency and power efficiency of factor five to ten. Across the individually optimized implementations, there are inter-architecture efficiency differences of typically one to two orders of magnitude. Moreover, maximum throughput performance and absolute power consumption both cover two orders of magnitude. Consequently, the choice of architecture for a particular application scenario depends on the importance attributed to the architectures' various properties.

The architecture-tailored parallelization concepts required for highest performance emphasize the mutual dependency between software and hardware the need for communication between both engineering domains.

Keywords — design space exploration, image processing architectures, stereo vision, semi-global matching

KURZFASSUNG

Die Einsatzmöglichkeiten von Bildverarbeitungssystemen zur maschinellen Umfeldinterpretation werden durch den rapiden technologischen und algorithmischen Fortschritt fortwährend erweitert. Jedoch erfordert der Entwurfsprozess solcher Systeme im zunehmenden Maße die Justierung einer Vielzahl an Designoptionen, die einen komplexen Einfluss auf die zum Teil gegenläufigen Entwurfskriterien wie z.B. Durchsatzrate, Verlustleitung, und Flexibilität haben. Die beständige Evaluierung aufkommender sowie bestehender Architekturen mit relevanten Bildverarbeitungsalgorithmen ist elementar, um frühzeitig die optimale Lösung bestimmen zu können. Dieses gilt insbesondere für Fahrerassistenzsysteme, bei denen für viele relevante Algorithmen noch keine schritthaltenden Implementierungen für eingebettete Systeme bekannt sind. Ein Algorithmus dieser Kategorie ist das Semi-Global Matching zur 3D-Umfelderfassung mittels passiver, kostengünstiger Stereokameras.

Die vorliegende Arbeit präsentiert Entwurf, Optimierung und Analyse von hocheffizienten Hardware- und Software-Implementierungen des Semi-Global-Matching-Algorithmus für ein diversitäres Set aktuell verbreiteter Architekturkonzepte. Die Evaluierung umfasst zwei ASPs (RISC und VLIW), eine GPU, einen Universalmehrkernprozessor, sowie eine dedizierte Architektur, die auf FPGAs und Standardzell-ASICs abgebildet wird.

Um eine einheitliche quantitative Bewertung der Architekturen hinsichtlich Siliziumfläche, Verlustleistungsaufnahme und Durchsatzrate durchführen zu können wird eine Entwurfsraumexplorationsumgebung eingeführt. Diese beinhaltet eine FPGA-basierte SoC-Entwicklungsumgebung, welche emulationsbasierte Analysen ermöglicht. Darauf basierend wird eine detaillierte Untersuchung der Implementierungsalternativen innerhalb einer Architektur (Intraarchitekturanalyse) sowie der Architekturen zueinander (Interarchitekturanalyse) vorgestellt. Die praktische Anwendung der SoC-Umgebung wird weiterhin durch den Aufbau eines echtzeitfähigen Stereovideosystems gezeigt.

Die Auswertung zeigt, dass ein Intraarchitekturoptimierungspotential von Faktor fünf bis zehn hinsichtlich Flächen- und Verlustleistungseffizienz besteht. Das Optimierungspotential zwischen den individuell optimierten Implementierungen der untersuchten Architekturkonzepte beträgt typischerweise ein bis zwei Größenordnungen. Weiterhin umspannen Durchsatzrate und absolute Verlustleistung jeweils einen Bereich von ca. zwei Größenordnungen. Konsequenterweise kann eine Festlegung der Architektur nur unter zusätzlicher Berücksichtigung der anwendungsspezifischen Bedeutung der jeweiligen Eigenschaften vorgenommen werden.

Schlagworte — Entwurfsraumexploration, Architekturen zur Bildverarbeitung, Stereobildverarbeitung, Semi-Global Matching

CONTENTS