

Kurzfassung

Mit der Einbeziehung psychoakustischer Modelle in technische Systeme für die Sprachverarbeitung, wie automatische Spracherkennung, Sprachgütebewertung in Kommunikationssystemen und digitale Hörgeräte kann eine signifikante Verbesserung der Verarbeitungsqualität erreicht werden. Die vorliegende Dissertation beschreibt sowohl eine Konzeption zur geeigneten Optimierung audiosignalverarbeitender Algorithmen, als auch einen Systementwurf für die Verifikation, den Test und den direkten Einsatz der dafür entwickelten hochintegrierten anwendungsspezifischen Schaltungen. Für die nichtlinearen adaptiven Nachregelschleifen des eingesetzten Oldenburger Perzeptionsmodells wurde mit Hilfe eines C/C++ Prototypen eine zweigeteilte Architektur mit Fest- und Fließkomma-Einheiten entwickelt, die in ihrer Präzision an die lokale Signaldynamik angepaßt worden ist (Operandenpräzision 14 Bit Mantisse für Addition, Subtraktion und Multiplikation und 6 Bit für Division bei jeweils 6 Bit exzeß-kodierter Charakteristik). Im Zusammenwirken mit einem dafür angepaßten Verfahren zur Sprachgütebeurteilung durch das Perzeptionsmodell selbst konnte die optimale Architektur bestimmt und validiert werden, indem nur die perzeptiv relevanten Signalverzerrungen am Ausgang des Perzeptionsmodells zur Berechnung eines Gütemaßes herangezogen wurden.

Für die Synthese der Algorithmen wurde kein konventioneller struktureller Ansatz gewählt. Vielmehr wurde ein durchgängiger Entwurfsfluß für die direkte Datenpfadsynthese mit einem speziellen CAD-Werkzeug (*Synopsys Behavioral Compiler*) entwickelt. Nach der Erarbeitung einer synthetisierbaren VHDL-Beschreibung des Algorithmus' wurden die mit dem C/C++ Prototypen spezifizierten VHDL-Fließkomma-Operatoren so erstellt, daß sie durch Operatorüberladung inferiert werden können. Dabei mußten insbesondere die Einflüsse und Rückwirkungen aus Architekturvarianten, Randbedingungen und Beschränkungen des Design-Systems, der untersuchten FPGA- (*Xilinx Virtex* und *Altera Flex*) und der ASIC-Technologien (*ES2-07*) auf die Optimierung des Entwurfsflusses und des Algorithmus' selbst berücksichtigt werden. Der Entwurfsfluß ist durch die Abläufe des abschnittsweise automatischen CAD-Werkzeugs nur grob vorgegeben und mußte durch eine skript-basierte Steuerung zum Erreichen der gewünschten Ergebnisse und zur Umgehung von Bibliotheks- und Werkzeuginkonsistenzen individuell angepaßt und optimiert werden. Das Verifikationsverfahren für Optimierungen am Algorithmus und dem Datenformat wurde so ausgebaut, daß die Simulation mit dem ursprünglichen C-Kode, dem optimal quantisierenden C++ Prototypen und den VHDL-Modellen in jeweils derselben Umgebung möglich ist. Das hier erarbeitete Hardware-Design konnte auf Verhaltens-, Register-Transfer- und Layoutebene korrekt simuliert werden; es benötigt in einem *Xilinx Virtex*-FPGA *XCV800* nur 19 % der Gatterressourcen (äquivalente Gatterzahl: 80.000) bzw. 10 % des Onchip-RAM-Speichers und erzielt eine Taktfrequenz von 40 MHz. Mit einer Gesamtverzögerung von 3,5 µs je Abtastwert werden die für eine Echtzeit-Datenverarbeitung maximal zulässigen 4,096 µs deutlich unterschritten. Mit der Integration aller Bestandteile des entwickelten Chipsatzes in ein kommerzielles Prototypensystem und der Einbindung in eine PC-basierte Arbeitsumgebung steht damit ein vollständiges und echtzeitfähiges Audiosignalverarbeitungssystem den gewünschten Anwendungen zur Verfügung.

Abstract

A significant processing quality improvement can be achieved for technical speech processing systems such as automatic speech recognition, speech quality estimation in communication systems and digital hearing aids by the integration of psychoacoustic models. The presented Ph. D. thesis describes a conception for a qualified optimization of audioprocessing algorithms as well as a system approach for verification, test and direct application of dedicated integrated circuits. A mixed fix point/floating point architecture was developed for the nonlinear adaptation loops of the implemented Oldenburger Perception Model. The resolution of these arithmetical units which can be scaled due to the local signal dynamic throughout the model (optimal operand precision 14 bit mantissa for addition, subtraction and multiplication and 6 bit for division each with a 6 bit excess-coded characteristic) has been specified using a C/C++ prototype. The optimized architecture has been estimated and validated in conjunction with an adapted method for the speech quality estimation by the perception model itself, since the perceptual relevant signal distortions at the model output are used for the calculation of a quality measure.

A comprehensive design flow for the direct datapath synthesis was chosen for the algorithms using a special CAD tool (*Synopsys Behavioral Compiler*) rather than a conventional structural synthesis approach. After establishing a synthesizable VHDL description of the algorithm itself the arithmetical VHDL operators specified with the C/C++ prototype have been developed for direct operator overloading and inference. Especially the influence and feedback of architectural options, boundary conditions, and limitations of the design tool and the tested FPGA (*Xilinx Virtex, Altera Flex*) and ASIC (*ES2-07*) technologies had to be taken into consideration for the optimization of the design flow and the algorithm itself. The design flow is roughly given by the partially automatic design tool and had to be adapted by script based controlling in detail to achieve the desired results and to avoid library and tool inconsistencies. The verification method for the algorithm and data type optimizations were extended for the simulation and validation of the original C-code, the optimally quantizing C/C++ prototype and the VHDL-models in the same environment. The presented hardware design has been correctly simulated on the behavioral, register-transfer and layout level. The overall usage of a *Xilinx Virtex* FPGA *XCV 800* is about 19 % for logic resources (equivalent gate count 80.000), 10 % for the internal RAM cells and the maximum clock rate is about 40 MHz. With an overall delay of 3,5 μs for the processing of one sample the real-time demands of at minimum 4,096 μs were clearly met. A real-time capable audio-signal processing system is available for the desired applications by the integration of all components of the developed chip set in a commercial hardware prototyping board and the implementation in a PC-based working environment.