

Abstract

When humans engage in spoken interaction, they assess the obvious features of the interaction partner, such as the verbal and non-verbal signals of the voice, as these contribute to the subjective evaluation of the interlocutor. In this book, the background, state of research, and own contributions to the assessment and prediction of talker quality that is constituted in voice perception and in dialog are presented. Starting from theories and empirical findings from human interaction major results and approaches are transferred to the domain of human-computer interaction. The main subject of this book is to contribute to the evaluation of spoken interaction in both, humans and between human and computer, and in particular to the quality subsequently attributed to the speaking system or person, based on the listening and interactive experience.

The theories, methods, and results presented are focusing on the first impression of the people engaged in such a vocal conversation. This means hearing a voice for the first time in either a passive scenario (listening only), or interacting with a person or computer for the first time (interactive scenario). By using the term "first impression", the research focus is set to the human perception and evaluation of voices and conversations experienced, which represents the beginning attitude formation of the participating human towards the (other) speaker, may it be a real person or a computer agent.

The main scientific contribution is not to psychological theory development, but it is an informed engineering approach for describing subjective quality as experienced by users. As mentioned in the subtitle, the major results of this book are the development of quantitative models of user ratings to represent subjective quality. The most important part of this modeling is the identification of relevant parameters as predictors. Such predictors of talker quality can be acoustic features, for the case of voices' quality that is typically assessed in passive scenarios, or data describing interaction behavior for the course of conversation. By modeling voice-stimulated user ratings, this book also contributes to the identification of the most important perceptual dimensions – as these factors provide the valuable insight in search for quantitative predictors, and derives basic insights and principles in designing and evaluating modern spoken conversational systems, with the aim of increasing and ensuring quality.

Zusammenfassung

In gesprochener Interaktion schätzen Menschen die offenbaren Merkmale des Interaktionspartners, wie etwa verbale und non-verbale Signale der Stimme, ein, da diese zur subjektiven Evaluation des Gesprächspartners beitragen. In diesem Buch werden Grundlagen, Forschungsstand und eigene Beiträge zur Feststellung und Vorhersage von Talker Quality, die in Stimmwahrnehmung und im Dialog erzeugt wird, dargestellt. Beginnend bei Theorien und empirischen Befunden zur menschlichen Interaktion werden Hauptergebnisse und Ansätze in die Domäne der Mensch-Computer-Interaktion überführt. Der zentrale Beitrag dieses Buches liegt in der Evaluierung gesprochener Interaktion zwischen Menschen und Mensch mit Computer und insbesondere in der Qualität, die sprechenden Systemen oder Personen auf Basis von Hör- und interaktiven Erfahrungen zugeschrieben wird.

Die Theorien, Methoden und Ergebnisse fokussieren sich auf den ersten Eindruck von Personen, die in solchen gesprochenen Konversationen beteiligt sind. Das bedeutet das erstmalige Hören einer Stimme, sei es einer passiven Szenarios (listening only), oder in Interaktion mit einer anderen Person oder einem Computer (interactive Scenario). Mit der Verwendung des Begriffes „first impression“ wird der Fokus der Forschung auf die menschliche Wahrnehmung und Evaluierung von Stimmen und das Erleben von Gesprächen gesetzt. Erste Eindrücke repräsentieren den Beginn einer Einstellungsbildung zu dem (anderen) Sprecher, sei es eine echte Person oder ein Computer.

Der wissenschaftliche Beitrag liegt nicht in psychologischen Theorienentwicklung, sondern es ist ein fundierter Ingenieursansatz zur Beschreibung subjektiver Qualität, wie sie von Nutzern erlebt wird. Wie im Untertitel genannt, bilden die Hauptergebnisse die Entwicklung von quantitativer Modelle von Nutzerurteilen subjektiver Qualität. Der wichtigste Teil dieser Modellbildung ist die Identifizierung relevanter Faktoren als Modellprädiktoren. Solche Prädiktoren von Talker Quality können akustische Merkmale sein, für den Fall von stimmlicher Qualität, die üblicherweise in passiven Szenarios erfasst wird, oder parametrische Beschreibungen des Interaktionsverlaufes. Mit der Modellierung stimm-stimulierter Nutzerurteile trägt dieses Buch auch zur Identifikation der bedeutendsten Wahrnehmungsdimensionen bei – da diese Faktoren wertvolle Einsicht für die Suche nach quantitativen Prädiktoren bieten. Außerdem werden grundlegende Erkenntnisse und Prinzipien für das Design und die Evaluierung moderner Konversationssysteme abgeleitet, um die die subjektive Qualität sicherzustellen und zu erhöhen.