Accepted Manuscript

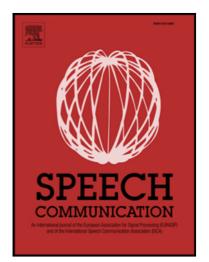
Speaker recognition from whispered speech: A tutorial survey and an application of time-varying linear prediction

Ville Vestman, Dhananjaya Gowda, Md Sahidullah, Paavo Alku, Tomi Kinnunen

 PII:
 S0167-6393(17)30263-7

 DOI:
 10.1016/j.specom.2018.02.009

 Reference:
 SPECOM 2541



To appear in: Speech Communication

Received date:6 July 2017Revised date:22 December 2017Accepted date:13 February 2018

Please cite this article as: Ville Vestman, Dhananjaya Gowda, Md Sahidullah, Paavo Alku, Tomi Kinnunen, Speaker recognition from whispered speech: A tutorial survey and an application of time-varying linear prediction, *Speech Communication* (2018), doi: 10.1016/j.specom.2018.02.009

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Speaker recognition from whispered speech: A tutorial survey and an application of time-varying linear prediction *

Ville Vestman¹, Dhananjaya Gowda², Md Sahidullah¹, Paavo Alku³, Tomi Kinnunen¹

¹School of Computing, University of Eastern Finland, Finland ²DMC R&D Center, Samsung Electronics, Seoul, Korea ³Department of Signal Processing and Acoustics, Aalto University, Finland

Abstract

From the available biometric technologies, automatic speaker recognition is one of the most convenient and accessible ones due to abundance of mobile devices equipped with a microphone, allowing users to be authenticated across multiple environments and devices. Speaker recognition also finds use in forensics and surveillance. Due to the acoustic mismatch induced by varied environments and devices of the same speaker, leading to increased number of identification errors, much of the research focuses on compensating for such technology-induced variations, especially using machine learning at the statistical back-end. Another much less studied but at least as detrimental source of acoustic variation, however, arises from mismatched speaking styles induced by the speaker, leading to a substantial performance drop in recognition accuracy. This is a major problem especially in forensics where perpetrators may purposefully disguise their identity by varying their speaking style. We focus on one of the most commonly used ways of disguising one's speaker identity, namely, whispering. We approach the problem of normal-whisper acoustic mismatch compensation from the viewpoint of robust feature extraction. Since whispered speech is intelligible, yet a low-intensity signal and therefore prone to extrinsic distortions, we take advantage of robust, long-term speech analysis methods that utilize slow articulatory movements in speech production. In specific, we address the problem

Preprint submitted to Speech Communication

March 7, 2018

^{*}This work contains limited portions of [1].

دريافت فورى 🛶 متن كامل مقاله

- امکان دانلود نسخه تمام متن مقالات انگلیسی
 امکان دانلود نسخه ترجمه شده مقالات
 پذیرش سفارش ترجمه تخصصی
 امکان جستجو در آرشیو جامعی از صدها موضوع و هزاران مقاله
 امکان دانلود رایگان ۲ صفحه اول هر مقاله
 امکان پرداخت اینترنتی با کلیه کارت های عضو شتاب
 دانلود فوری مقاله پس از پرداخت آنلاین
 پشتیبانی کامل خرید با بهره مندی از سیستم هوشمند رهگیری سفارشات
- ISIArticles مرجع مقالات تخصصی ایران