International Journal of Innovative
Computing, Information and Control
Volume **7**, Number **8**, August **2011**

ICIC International ©2011 ISSN 1349-4198
pp. **4929–4940**

BODY-CONDUCTED SPEECH RECOGNITION IN SPEECH SUPPORT SYSTEM FOR DISORDERS

Shunsuke Ishimitsu[1], Kouhei Oda[1] and Masashi Nakayama[2]

[1]Graduate School of Information Sciences
Hiroshima City University
3-4-1 Ozuka-Higashi, Asa-Minami-Ku, Hiroshima 7313194, Japan
ishimitu@hiroshima-cu.ac.jp

[2]Department of Electrical and Computer Engineering
Kagawa National College of Technology
355 Chokushi, Takamatsu, Kagawa 7618058, Japan
m-nakayama@t.kagawa-nct.ac.jp

Received May 2010; revised October 2010

Abstract. *People with speech disabilities often face communication problems in daily conversation. Substitutes for speech can be used in communication, but these are typically not used sufficiently frequently in daily conversation to be readily understood. We developed a speech support system using body-conducted speech recognition. This system retrieves speech via a transfer function from body-conducted speech, using a recognition method to determine sub-word sequences and duration time. To construct this system, it was necessary to first examine the effectiveness of body-conducted speech recognition in speech disorders. As a first step, we investigated continuous word unit speech recognition using an acoustic model that was not adapted to body-conducted speech for patients with speech disorders. We analyzed each parameter of this speech, and performed body-conducted speech recognition. We concluded that an adaptation method using body-conducted speech recognition is a valid technique for achieving high recognition performance in speech disorders.*

Keywords: Cancer of the pharynx, Substitution speech, Esophageal speech, Body-conducted speech, Speech recognition, Transfer function, Cross-spectrum method

1. **Introduction.** Pharyngeal cancer is the cause of many disorders and is associated with an increasing rate of pharyngeal surgery [1]. Although most patients recover quickly after surgery, speech disabilities commonly occur after the removal of the pharynx. As a result, patients typically communicate with one of two forms of substitute speech, esophageal vocalization or speech reconstruction, which is chosen according to their condition. Esophageal speech is inexpensive and does not require surgery for vocalization. Although speech reconstruction requires surgery, patients with most types of speech disorder can easily perform speech communication after the procedure. However, substitute speech does not provide the optimal fundamental frequency, high frequency component or power required in daily conversation. Thus, patients with speech disorders commonly face communication problems in daily life.

To resolve these problems, researchers have attempted to improve the sound quality of substitute speech [2-9]. Akimoto et al. improved the quality of retrieval using a fundamental frequency [2] and Mitsuyoshi et al. proposed a human emotional analysis system for speech [3]. Nakamura et al. constructed a voice conversion system using transmitted artificial speech [4]. Ando et al. proposed a speech synthesis system for a Chinese language training system [5]. In addition, we previously proposed a speech