

SPEECH SUPPORT SYSTEM USING BODY-CONDUCTED SPEECH RECOGNITION FOR DISORDERS

MASASHI NAKAYAMA AND SHUNSUKE ISHIMITSU

Graduate School of Information Sciences
Hiroshima City University
3-3-1 Ozuka-Higashi, Asa-Minami-Ku, Hiroshima, 7313194, Japan
m-nakayama@kagawa-nct.ac.jp; ishimitu@hiroshima-cu.ac.jp

Received October 2008; revised February 2009

ABSTRACT. *The prevalence of several disorders, associated with cancer of the pharynx is increasing and many disorders require surgical removal of the pharynx. Although successful recovery has been reported after surgery, most people with disorders communicate by substitution speech. A solution that is commonly used to deal with such communication disorders is esophageal speech. However, esophageal speech sound is not a clear type of speech and does not have a sufficient volume or frequency for use in daily life. Therefore, we investigated development of a speech support system using body-conducted speech recognition (BCS) for disorders. This system uses two features: one is body-conducted speech recognition and the other is a sub-word unit transfer function database. First, a disorder utters esophageal speech into the system. The system then estimates a sub-word sequence and its duration time using body-conducted speech recognition. The system retrieves the speech using a sub-word unit sequence information with transfer functions estimated between speech and body-conducted speech. Since a speaker's own speech and their own body-conducted speech are used in this system, it is possible to change into the speaker's own speech from body-conducted speech for disorders. As the first step, we verified the effectiveness of the proposed method using body-conducted speech under calm conditions such as a quiet room.*

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

1. **Introduction.** In recent years, the number of people with disabilities that impede normal speech communication has increased. Pharyngeal cancer is the cause of many disorders and is associated with an increasing rate of pharynx-related surgery [1]. Although most affected patients recover well after surgery, they often become speech disorders. Therefore, these people have to overcome speech communication problems in daily conversations.

The most common solution used for speech disorders is esophageal vocalization, which is inexpensive and does not require surgery for vocalization. Esophageal vocalization is explained as follows. First, a speech disorder inhales air into the stomach and then breathes out into the air. The new glottis in the lower pharyngeal mucous membrane then vibrates, changing air into esophageal speech through the articulation organ between the pharynx and mouth. In this way, a speech disorder can generate esophageal speech. However, esophageal speech does not provide an optimal fundamental frequency, high-frequency component, or power for daily conversations. Therefore, speech disorders result in problems of communication in noisy environments encountered in daily life.

To solve these problems, many researchers have attempted to improve the sound quality of esophageal speech. The following studies were carried out to solve these problems.