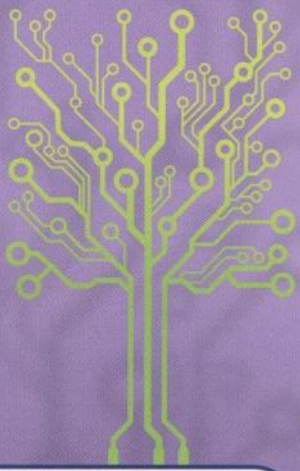

بیست و یکمین کنفرانس مهندسی برق ایران

دانشگاه فردوسی مشهد ● ۲۴ تا ۲۶ اردیبهشت ۱۳۹۲

21st Iranian Conference on Electrical Engineering

14 - 16 May 2013 ● Ferdowsi Univ. of Mashad

باسمه تعالی

بدین وسیله گواهی می شود که مقاله زیر در بیست و یکمین کنفرانس مهندسی برق ایران ارائه شده است.

"Computation Saving in a SRP-PHAT Sound Source Locator Variant"

Authors: Mohammad Seifipour, Saeed Seyedtabaii

میر مجتبی میرصالحی

دبیر کمیته علمی کنفرانس

Computation Saving in a SRP-PHAT Sound Source Locator Variant

M. Seifipour *, S. Seyedtabaii **
* Elec. Eng. Department, Shahed University , mohammadseyfipor@gmail.com
** Elec. Eng. Department, Shahed University, **stabaii@shahed.ac.ir**

Abstract: *The well-known Steered Response Power–Phase Transform (SRP-PHAT) speech source locator enjoys robustness and renders accurate results in high reverberation and medium noise situations. However, it suffers from high computation burden, which debars it from many applications. Various arrangements have been developed during the recent years to lessen the calculations and facilitate its implementation. In this paper, a new hybrid scheme is suggested and compared its performance and costs with several well-known contestant algorithms. The simulation results confirm that the algorithm performance is highly promising and noticeable.*

Keywords: Sound localization, Microphone array, SRP-PHAT.

1. Introduction

One of the main complications in video conferencing, remote security surveillance, advanced human–computer interaction, gunshot positioning, video-gaming and robotic applications is the accurate determination of sound source location. In video conferencing, by localizing the sound source, it is possible to aim a pan-tilt-zoom camera to the estimated position for capturing a high-resolution image or face to be monitored and or picking high quality speech by microphone array beam forming.

Sound source locators are loosely divided into three categories. 1) Time difference of arrival (TDOA): an indirect two-step approach that first, the TDOA between microphone pairs are estimated and then, the estimate of the source position based on the geometry of the microphone array and the estimated delays is optimally computed. The two-stage algorithms are fast, however, in noisy environments as the estimation of TDOA deteriorates, the second-stage search fails in rendering accurate results [1, 2]. TDOA based methods are the most commonly used ones in practice 2) Spectral estimation: It is a one stage high-resolution spectral estimation method [3, 4], which also handles multiple-source cases [3]. 3) SRP family: It is a one stage steered beamformer, which its idea is rooted in antenna array design & processing for RADAR. In the Steered Response Power–Phase Transform (SRP-PHAT) algorithm, a candidate source maximizing the output of a steered delay-and-sum beamformer is sought.

Generally, the steered-beamformer strategies are recommended for critical applications where robustness is important [3]. The method can accurately find the sound source location in low noise environments, under relatively heavy reverberations [5]. The algorithm drawback, however, is high computation costs [5, 2]. Taking into account the main requirements of acoustic localization and tracking algorithms where robustness to acoustic disturbances and low computation are immediate requirements, the algorithm fails in supporting the latter.

In this respect, an improved SRP-PHAT algorithm based on principal eigenvector has been suggested in [6] where sound source location is estimated from the principal eigenvector computed from the frequency-domain correlation matrix. In another attempt, in [7], stochastic region contraction (SRC) is proposed to combat the computation cost problem. A saving of nearly three orders of magnitude was achieved. No degradation in location-estimation performance was observed for "four" different locations and different SNR's versus original SRP-PHAT. Coarse-to-fine region contraction (CFRC), also based on the concept of region contraction, is introduced in [8]. It is shown that CFRC has some computational advantages while the performance and the average computational saving achieved by SRC are still maintained. An improved SRP–PHAT is presented in [9] that it reduces a two-dimension searching space into a couple of one-dimension spaces by using an orthogonal linear array.

A different strategy has been put forward in [10] where, instead of evaluating the SRP function at fine spatial grid nodes, the surrounding volume around nodes of a coarse spatial grid is searched by the Generalized Cross Correlation (GCC) lag space corresponding to the volume surrounding each point of a coarse grid. By this technique, it is claimed that while the computation is highly lowered, the accuracy remains as if a relatively fine grid has been adopted.

In this paper, a new hybrid algorithm is presented. The