

# Two-to-Five Upmix Methods

Pantelis Papastergiou

January 2018

## Abstract

The aim of several 3D audio concepts and products is to create a more immersive, engaging and natural-sounded listening experience. Emerging audio signal processing techniques, make it possible for regular stereo recordings to be compatible and reproduced with multichannel home theatre or automotive loudspeaker audio systems.

In this thesis, various existing methods are investigated and implemented for converting stereo recordings to four or five channels in the primary-ambience extraction (PAE) framework. In that, audio signals are often considered linear combinations of primary and ambient components. The former are assumed to be correlated, whereas the latter uncorrelated. The basic function of the upmix systems is to remove the correlated components from the electronic audio material, which are intended for playback with loudspeakers behind the listeners, in a 3/2 or 2/2 configuration. That way the decomposition facilitates the appropriate rendering for spatial enhancement.

The upmixers, either keep the initial stereo recording in the frontal loudspeakers or add a third central channel in the frontal setup to allow for off the "sweet spot" listening. All the methods are implemented in the frequency domain using the widely known short time Fourier transform (STFT) technique, except one. Central in the development of the algorithms in frequency domain are the method of Principal Components Analysis (PCA), the least squares estimates (LS), the normalized least mean squares (NLMS) adaptive filter and certain ambience masking functions. On the other hand, the core of the only time domain method is the least mean squares (LMS) adaptive filter.

Assessment of the new upmix systems was accomplished in an objective and subjective way; firstly, using performance measures such as the ambience energy fraction ( $E_A$ ) and the cross-correlation coefficient of primary and ambient components ( $\phi_P$  and  $\phi_A$  respectively), and secondly with a listening test which requires from the participants to judge the systems according to the overall impression.

The objective and subjective evaluation results suggest that a subjectively tuned ambience masking function and the frequency domain NLMS algorithm provide both promising upmix solutions and computational advantage.

Keywords: upmix, primary ambience extraction, adaptive filtering, spatial enhancement.