

PRINCIPAL PITCH EXTRACTION OF NON-SYMMETRIC VIBRATO TONES

Bruno A. Mesz, Manuel C. Eguia

Laboratorio de Acústica y Percepción Sonora. Universidad Nacional de Quilmes.

Vibrato is a performing technique widely used by singers in all cultures, consisting in a quasi-periodic frequency modulation. Although multiple frequencies are present in a vibrato note, when listening to a singer one tends to extract a single pitch, called *principal pitch*. It is still a matter of debate what is the principal pitch in vibratos with a nonsymmetric frequency profile (Shonle and Horan, 1980; D'Alessandro and Castellengo, 1994). This kind of vibrato is produced by some singing and instrumental techniques.

To study this issue, we made a pitch detection psychoacoustical experiment using vibratos with portions of fast and slow varying frequency. Although synthetically generated, examples of similar frequency profile could be found for instance in South Indian music (Krishnaswamy, 2003).

For the purpose of explaining the results of our experiment, we propose a biologically motivated model for principal pitch extraction based on computations that could be performed by the auditory system. We could determine a narrow range for the free parameters in our model to fit the data across all subjects, and moreover we could fit the data of other experiments on vibrato in the literature, with the same values of the parameters.

Our formulation is based in the idea, proposed in a recent paper (Gardner and Magnasco, 2006), that the auditory system may use a *reassigned* representation of sound stimuli in which time and frequency localization is computed using phases of oscillations. This representation maps frequencies physically present in sound to new values called *instantaneous frequencies*. This task could be implemented neurally using the time interval between action potentials in the nerve fibers. In this way an estimation of the instantaneous frequency is obtained for each auditory channel. If these estimations are very similar in a group of neighboring channels, we say that there exists a high degree of *consensus*, following the terminology of Magnasco and Gardner. We give a way of computing consensus locally in each auditory channel employing only phase information. The model then predicts principal pitch in vibrato as a weighted average of the instantaneous frequencies extracted, with a perceptual weight proportional to consensus and to a compressive function of channel amplitude.

References:

Shonle, J.I. and Horan, K.E.: "The pitch of vibrato tones", The Journal of the Acoustical Society of America—January 1980—Volume 67, Issue 1, pp. 246-252.

D'Alessandro, C. and Castellengo, M.: "The pitch of short-duration vibrato tones", The Journal of the Acoustical Society of America—March 1994—Volume 95, Issue 3, pp. 1617-1630

Krishnaswamy, A.: "Pitch measurements versus perception of South Indian classical music", Proceedings of the Stockholm Music Acoustics Conference (SMAC 03), Stockholm, Sweden.

Gardner, T. and Magnasco, M.: "Sparse time-frequency representations", PNAS, April 18, 2006, vol. 103, no 16, 6094-6099.