App development with audio applications from m-file to app
Project 1

Daniel Jankovic

June 28, 2015

Voice Activity Detector (VAD) is a technique used in signal processing to detect the presence of human voice in a signal. It is considered an important technology in speech based communication and today we have various types of applications that use it. Therefore various VAD algorithms have been developed to provide the needed features.

There are different kind of stand-alone commercial baby monitor on the market today. From the most basic, that use one-way radio communication, to advance two-way communication monitors that use signal processing to transmitt audio when a certain threashold of sound activity has been reached. Now days there are also baby video monitors that broadcast both audio and video when the sensors notice movement. Since the monitor system relies on radio signals to communicate between the units there is a risk that the signal strength will weaken or not reach the receiver becuase it needs to penetrate through multiple walls of varying thickness. The stand-alone monitor focuses on reliability but little is known about the security features. It is possible to assume that the communication is uncrypted, in some products, and therefore introduces a potential risk for intrusion of peoples privacy.

To resolve the issuses broughten up above, an application such as the BAD can can be made more portable, versatile and secure with the help of todays smartphone technology togehther with VAD.

Examples of VAD algorithms are short-time energy, where energy levels are calculated for each frame to detect voice, unvoiced or silenced regions. Voiced regions will have higher energy levels. Zero-crossing rate (ZCR), is the rate at which a signal changes from plus to minus and back. The higher the rate the higher the frequency which indicates voice activity. Linear predictive coding (LPC) is one of the most powerful speech analysis techniques which gives an extremely accurate estimate of the speech parameters. The algorithm is very complex but the basic

References

[1] R. Cohen, Y. Lavner, *Infant Cry Analysis and Detection*, 2012

[2] E. Verteletskaya, K. Sakhnov *Voice Activity Detection for Speech Enhancement Applications*, 2010

[3] N. Westerlund, M. Dahl *Speech Enhancement using an Adaptive Gain Equalizer* 2003