

CHAPTER 8

LIMITATIONS AND FUTURE SCOPE

Conclusion:

The work embodied in this thesis presents a speech and speaker authentication system with reference to the Bodo language, one of the most popular tribal language in North-East region. In this research work Artificial Neural Network (ANN) is used as the prime authenticator tool. The main objective of my work documented in presented thesis is to develop a BODO speech and speaker Authentication system, christened as SPEaker Recognition Identification and Authentication System form BODO, in short SPERIA B.

New insight gained from this research work are :

1. High level of authentication is achieved by multilevel feature vectors application.
2. Artificial Neural Network's effectiveness in Bodo speaker authentication is proved.
3. This work may play important role to preserve a language like Bodo having comparatively fewer users.
4. Multilayered and multiple feature vector application significantly enhances the authentication rate against the single vector based experiment.

Future trends and Prospect:

At the end we must accept that in spite of successful authentication rate, certain challenges remain unsolved. Even then my thesis will definitely help the future investigators; they may use it as the launching pad for further investigation.

Some of the areas identified for future work are:

a) Robustness:

This thesis is limited to a closed set of speakers and a closed set of words mainly, except in the last experiment environment where some impostors were experimented. But to design a robust authentication system, unknown speakers from unknown/robust environment have to consider. Present work can be a launching pad to extend the research work further for more effective and efficient robust speech/speaker authentication system in Bodo language.

b) Noise free Speech samples:

High Degree of Authentication accuracy achieved in SPERIA B is contributed by the fact that voice samples are recorded in noise free laboratory environment. Further silence/noise removal is performed. But in reality speaker will utter in robust situation in robust environment.

c) Small Database:

Our database consists of a relatively small number of speakers and words. More over Speech samples considered in this thesis are basically emphasized to be emotionless, in other words only general speaking style mode. In real environment, conversations will carry emotional quotient too, most of the time, in most of the cases.

d) The parameterizations or classifications are done independently. The combined results of these classifiers may produce a more authentic and realistic applications in any robust environment. This scope would enhance the current proposition more realistically.