

XVIII Encontro de Jovens Pesquisadores Universidade de Caxias do Sul - 2010

PASSVOX - Sistema para Coleta de Áudio Tolerante à Falha

Diego Silva do Espirito Santo (BIC/UCS), Andre Gustavo Adami (Orientador(a))

Atualmente, a administração da informação se tornou uma temática comum, sobretudo devido a sua importância nos meios de comunicação. Em tal meio onde a informação é falada e dinâmica e tal locução representa uma possível informação relevante, o armazenamento e tratamento das informações têm papel fundamental. Nas informações armazenadas em áudio, as características podem ser extraídas para se identificar quem fala, o que fala ou mesmo se fala. Como uma solução para o tratamento das informações geradas através de conversas gravadas, o projeto PassVox visa o desenvolvimento de um framework para reconhecimento automático de locutor para bases de áudio. Este framework será composto por algoritmos de detecção de atividade de fala, segmentação de locutor, reconhecimento de locutor e outros algoritmos necessários para o processamento do áudio. Mas para que os algoritmos citados identifiquem ou detecte a fala ativa de uma pessoa, é necessário que exista informação disponível para tal – Fonte de áudio. Neste trabalho descrevemos um programa gravador de voz, chamado de WaveRecorder, que compõe o framework do PassVox. WaveRecorder é um software para captura e armazenamento de áudio e foi projetado, em c#, para ser executado sob o sistema operacional Windows, com a plataforma .Net instalada. Suas funcionalidades principais são o armazenamento do áudio em arquivo e a segurança de que os dados coletados não serão perdidos caso alguma falha no sistema aconteça. Ele armazena em disco o som capturado através da placa multimídia do computador a cada período pré-determinado, permitindo a gravação parcial de um arquivo de áudio (o que não acontece em sistemas convencionais de gravação de áudio). Usando classes do dotNet e API do Windows, o gravador é simples e fácil de manipular e permite a definição da taxa de amostragem e o número de canais. Foram realizados dois testes específicos para garantir a integridade do áudio capturado, com perda limitada de 1 minuto. No primeiro teste, a aplicação foi encerrada abruptamente para verificar a perda de dados de áudio. O resultado foi a perda de um segmento de áudio menor do que 1 minuto, conforme especificado. No segundo teste, a aplicação foi executada por um longo tempo a fim de testar a capacidade de lidar com grandes quantidades de dados. A execução ficou ativa por mais de 24h, atendendo as especificações definidas. A aplicação foi testada com rádios comunicadores para simular a conversa entre duas pessoas distintas.

Palavras-chave: coleta de áudio, processamento de voz, rádios comunicadores.

Apoio: UCS