

INTERFACE COMPUTACIONAL DE BIOMETRIA COMO FERRAMENTA DE APOIO À PERÍCIA DE CONFRONTO DE VOZ

Gustavo Moreira da SILVA¹
Ângela Abreu Rosa de SÁ¹
Vinícius Naves Rezende FARIA¹
Alcimar Barbosa SOARES¹
Issa Khalil Georges DIB NETO¹
Hélio PEIXOTO²

- RESUMO: A Engenharia Biomédica Forense utiliza as atribuições da Engenharia Biomédica para fins Jurídicos, englobando questões de identificação humana abordadas pela biometria. A biometria é uma ciência que promove processos de identificação de indivíduos baseada em características comportamentais ou fisiológicas, apresentando diversas aplicações, dentre elas a elucidação de crimes. Dentre os vários tipos de biometria existentes, a biometria de voz consiste em um processo de identificação que avalia a fala do indivíduo, bem como suas características comportamentais durante a fala. Assim, esse trabalho propõe o desenvolvimento de uma ferramenta computacional que integra vários parâmetros acústicos da fala, numa interface simples, com o objetivo de apoiar a perícia de reconhecimento de locutor. Os resultados apresentados inferem que a ferramenta pode ser eficiente no auxílio ao perito, pois apresenta, automaticamente, uma comparação entre os parâmetros acústicos dos áudios analisados; o que facilita e agiliza a emissão do laudo pericial.
- PALAVRAS-CHAVE: Biometria; engenharia biomédica forense; fonética forense; identificação de locutor.

1 Introdução

Os aspectos comuns de comportamento dos indivíduos expressam características adquiridas do grupo regional e social em que tiveram a sua formação cultural. Assim, os traços linguísticos de um falante também são marcados pelo ambiente de sua formação e crescimento. Portanto, ao se comunicar através da fala, o indivíduo expressa o seu pensamento de acordo com a sistemática de suas ideias, do seu conhecimento, dos padrões de comportamento adquiridos, e de atitudes que fornecem pistas do seu perfil e ambiente de origem (BRAID, 2004). Assim, a capacidade humana para reconhecer indivíduos apenas pela voz pode se manifestar cotidianamente nas mais variadas situações, inclusive

¹ Universidade Federal de Uberlândia – UFU, Faculdade de Engenharia Elétrica, Laboratório de Engenharia Biomédica–BioLab, CEP: 38400-000, Uberlândia, MG, Brasil. E-mail: angela@eletrica.ufu.br

² Perícia Criminal, Departamento de Polícia Federal, Brasília, DF, Brasil.

sob condições pouco favoráveis, envolvendo diferentes tipos de distorções e limitações do sinal acústico (FIGUEIREDO, 1994)

Nessa perspectiva, a fala não transmite apenas uma mensagem em sentido estrito, como numa simples frase escrita numa folha de papel. Ela veicula, ao mesmo tempo, informações de caráter pessoal que qualificam o falante quanto à sua formação linguística e cultural, sua origem regional e social, além de estabelecer dados sobre sua identidade física e o seu estado de humor momentâneo (BRAID, 2004). Esses fatores descrevem a voz como um exemplo de identificação pessoal biométrica. A identificação humana pode servir de instrumento para diversas ciências, dentre elas, as Ciências Forenses ou a Criminalística.

As atribuições da Criminalística envolvem várias ciências, dentre elas a Engenharia Biomédica, que é o ramo das ciências da engenharia que tem como objetivo o desenvolvimento e a aplicação de tecnologias voltadas para a área da saúde; incluindo o estudo dos sinais biológicos (eletromiográfico, biométricos, impedância da pele, eletroencefalográfico, eletrocardiográfico e voz) para a identificação humana e auxílio no diagnóstico de patologias e avaliação psicológica. No âmbito das Ciências Forenses, a Engenharia Biomédica Forense, utiliza estas atribuições da Engenharia Biomédica, dentre elas a identificação biométrica, para fins jurídicos.

Define-se Biometria como a ciência que busca o reconhecimento automático de indivíduos por meio de métodos estatísticos e processamento de sinais que mapeiam elementos anatômicos ou fisiológicos dos seres humanos – por exemplo, íris, voz, face, movimento labial, batimentos cardíacos, impressões digitais ou palmares, geometria da mão, forma de deslocamento e DNA - em características numéricas (BESACIER, 2003; MASON, 2005). Assim, a biometria corresponde a processos de identificação de indivíduos, baseados em suas características comportamentais ou fisiológicas, com aplicações no controle de acesso a redes ou na elucidação de crimes.

Dentre os vários tipos de biometria existentes, a biometria de voz consiste em um processo de identificação através da verificação de características comportamentais de um indivíduo, avaliando a elocução do mesmo, comparando as amostras do sinal de voz. Atualmente, utilizando diversas metodologias, é possível o perito realizar um confronto de dois áudios e confirmar se tais áudios foram produzidos ou não pelo mesmo aparelho fonador. Antigamente, a identificação de locutores se pautava no juízo de valor de determinadas pessoas, que utilizavam da análise qualitativa para este propósito. Com o desenvolvimento científico da área forense, desenvolveu-se melhores metodologias de análise envolvendo estudos linguísticos e análise computacional, sendo esta possível com o avanço da tecnologia da informação.

No entanto, a identificação forense de locutores apresenta deficiências no que envolve a automatização da metodologia utilizada pelos peritos da área. Tais deficiências estão relacionadas com a ausência de um sistema computacional que seja capaz de calcular as características de interesse no processo de identificação, agregando-as em única rotina de análise. Atualmente, os métodos computacionais que são utilizados para processar o sinal de voz, são constituídos por uma gama de softwares (BOERSMA e WEENINK, 2012; FONOVUEW, 2012; VOXMETRIA, 2012).

Neste contexto, observa-se que essas atuais ferramentas computacionais de análise vocal não realizam o confronto automático da voz e não são dedicadas exclusivamente para a análise pericial. Isso ocorre porque esses sistemas não foram desenvolvidos tendo

como fim a atividade pericial. Portanto, elas não têm como foco principal a identificação de locutor propriamente dita, e sim, somente a análise acústica do sinal da fala.

Face ao exposto, o objetivo deste artigo é apresentar a elaboração de um sistema de processamento digital de sinal de voz capaz de fornecer suporte à perícia de confronto de voz. Este sistema tem como objetivo colaborar com o trabalho pericial, apresentando uma interface amigável ao contexto da perícia de voz e irá, automaticamente, realizar a extração das principais características acústicas da voz, apresentando-as de forma comparada. Assim, esta ferramenta poderá auxiliar o perito na avaliação das convergências e divergências de cada parâmetro e, desta forma, colaborar para a conclusão do laudo pericial.

2 Identificação forense da voz

Fonética (do grego “*phonetiké*”) é uma ciência que estuda a física relativa aos sons da fala humana, observando propriedades físicas dos sons da fala, e os processos fisiológicos da sua produção (BORBA, 1975). Assim, a Fonética Forense é a utilização da Fonética para os propósitos forenses ou jurídicos.

Para a prática forense utiliza-se a terminologia identificação forense da voz. Tal termo refere-se a um procedimento de fundamental importância em situações em que a única forma de se atribuir a autoria de um ilícito penal, ou ainda desvincular uma pessoa dele, é determinar a autoria da fala em mídia (SCATENA, 2010). A identificação forense da voz exige do perito um prévio conhecimento sobre a anatomia do sistema respiratório, a fisiologia de produção da fala, conceitos sobre linguística. Todos esses conceitos integrados complementam um estudo sobre fonética forense, que é o estudo das propriedades biométricas da voz humana para fins jurídicos.

O processo de identificação de locutor é realizado pelos peritos. Os peritos são os profissionais das Ciências Forenses, eles detêm conhecimento técnico e científico e são capazes de esclarecer fatos de interesse judicial, e no exercício das suas atribuições utilizam da criminalística (SCATENA, 2010; VELHO *et al.*, 2012). Neste sentido, os peritos em fonética forense atuam trabalhando com as evidências de voz a serem examinadas.

O Perito Criminal Antonio César Morant Braid (BRAID, 2004) explicita: “*A Fonética é um extraordinário meio que pode ser utilizado para identificação humana, pois na fala estão contidos traços característicos de um indivíduo, da sua origem regional e social, do seu estado emocional momentâneo e outras informações que podem ser inferidas a partir do material de fala. A utilização da Fonética na área forense é necessária e, quase sempre, insubstituível para a solução de crimes em relação aos quais existam vozes registradas em algum tipo de mídia*”.

Na identificação do locutor são considerados dois importantes tipos de análises: a análise perceptual e a análise Acústica (BEHLAU, 1993). Nas fases iniciais de identificação de locutor, a análise perceptual tem a função de determinar qualitativamente os traços fonéticos de cada indivíduo, fazendo a extração dos aspectos da fala para comparação perceptual com os traços de outro locutor. Este método corresponde a exames qualitativos que abrangem a Fonética Articulatória da produção dos sons da fala, Fonologia, e a Linguística (WOLF, 1972; MORISSON, 2004; FLORES, 2008; BEIGI, 2011). Já a análise acústica permite calcular as grandezas físicas que são produzidas a

partir de configurações específicas do aparelho fonador, e apresentar graficamente sua evolução no tempo. Esta mensuração é realizada em segmentos e em supras segmentos da fala (BRAID, 1999;BRAID, 2004).

A identificação de locutores partiu de um método que se baseava única e exclusivamente na audição humana. Tal método foi evoluindo e a análise perceptual foi ganhando métodos mais eficientes, como um maior estudo da linguística e a utilização de conhecimentos da fonética articulatória (WOLF, 1972; RABINER e SCHAFER, 1978). Com o advento da ciência do processamento digital de sinais, vários tipos de sistemas e técnicas foram sendo desenvolvidos ao longo do tempo com o objetivo de melhorar a acurácia da identificação dos locutores (TOSI, 1979; RABINER e JUANG, 1993).

Baseado nas análises perceptual e acústica, os laudos periciais são confeccionados a partir das análises das características extraídas e classificadas em marcadores como aquele proposto por (TOSI, 1979). Esses marcadores que contém diversos aspectos da fala quando condensados permitem ao perito tomar a decisão de identificação ou eliminação do conteúdo analisado.

Atualmente, não existe nenhum sistema computacional que auxilie, de forma rápida e precisa a decisão de um laudo pericial. O perito precisa utilizar vários sistemas para extrair as informações vocálicas necessárias e fazer, manualmente, uma tabela com as principais informações quantitativas para realizar uma comparação acústica. Este procedimento requer muito tempo e depende muito da habilidade/experiência do perito.

E é baseado nessas deficiências de ferramentas para auxiliar o trabalho do perito no confronto de voz que está a proposta desse trabalho: o desenvolvimento de um sistema computacional de fácil manipulação e que faça a extração, de forma automática, de diversos parâmetros acústicos de dois sinais de voz. E ainda, o sistema deve fazer uma comparação estatística (margem de erro) entre esses parâmetros, auxiliando, assim, o perito na decisão do confronto das vozes. Estas propostas permitirão mais agilidade e rapidez para a análise acústica em uma perícia.

3 Metodologia

O software desenvolvido foi denominado *VoiceAnalisys*. Com base nos atuais procedimentos adotados nas perícias de reconhecimento de locutor, este sistema propõe ser uma ferramenta de apoio ao perito. Dessa forma, o projeto do sistema foi desenvolvido com o objetivo de permitir ao perito realizar uma análise acústica comparada entre dois áudios, numa interface intuitiva, simples e amigável.

3.1 As fases do projeto

A Figura 1 apresenta as fases de desenvolvimento do projeto: (a) Iniciação, (b) Planejamento (c) Execução e (d) Encerramento.

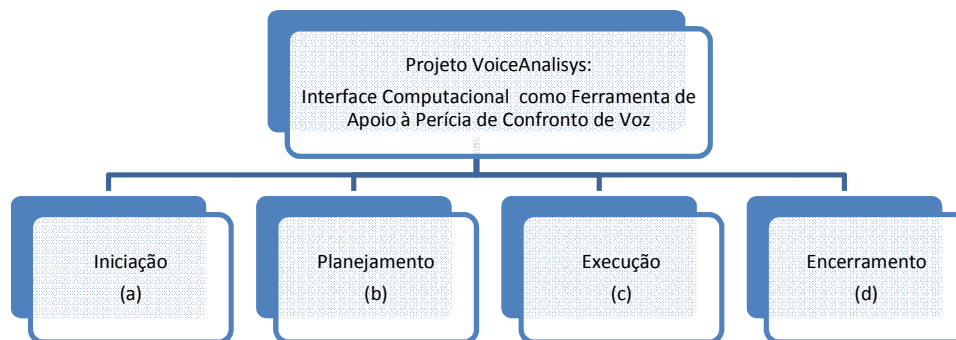


Figura 1 - Etapas de desenvolvimento do projeto “VoiceAnalysys”.

(a) Iniciação

A iniciação do projeto constituiu-se na fase marcada pela concepção e pela definição inicial do escopo do projeto. Além disso, foi realizada uma extensa pesquisa bibliográfica a fim de possibilitar a definição dos requisitos do sistema, contextualização da realidade da prática das perícias de voz, definindo as principais técnicas de processamento digital de sinais de voz que são utilizadas no confronto de voz.

(b) Planejamento

Após a pesquisa bibliográfica foi possível observar nos softwares existentes os requisitos que poderiam agilizar e facilitar o trabalho do perito. Esta etapa permitiu a concepção da proposta de um sistema computacional para apoio na perícia de identificação de locutor.

Elaborada a proposta de trabalho, o passo seguinte é o levantamento de requisitos que é referente ao processo de compreensão do problema aplicado ao desenvolvimento do software. Com base nos softwares similares estudados, e conforme a proposta estabelecida nesse trabalho levantou-se os seguintes requisitos:

(c) Requisitos Funcionais (RF)

RF_01: O software deve apresentar, de forma comparada, os seguintes gráficos:

- Transformada Rápida de Fourier (FFT);
- Intensidade (*loudness*);
- Transformada Rápida de Fourier de Curto Prazo (STFT);
- Conversão para a Escala Mel;
- Histograma do Pitch;
- LPC (Envoltória da FFT);
- Densidade de energia;

RF_02: O software deve realizar uma comparação dos seguintes parâmetros acústicos:

- Frequência Fundamental (*pitch*);
- Intensidade (*loudness*);
- Formantes;
- Jitter;
- Shimmer.

RF_03: O software deverá realizar análises estatísticas dos gráficos de densidade de energia e envoltória da FFT.

RF_04: O software deverá reproduzir os áudios analisados.

RF_05: O software deverá exportar um relatório formato texto com os parâmetros acústicos do áudio 1 e do áudio 2.

(d) Requisitos Não Funcionais (RNF)

Os Requisitos não funcionais descrevem as qualidades do sistema. A qualidade reflete diretamente a satisfação do usuário e envolvidos com o sistema. Assim, os requisitos não funcionais do “VoiceAnalysys” são:

RNF_01: O sistema deverá ser compatível com diversas versões do sistema operacional (XP, Vista, 7 e 8)

Justificativa: Por ser um sistema operacional mais comum entre a população em geral, inclusive nos computadores onde são realizadas as perícias.

RNF_02: O sistema deverá ser desenvolvido na linguagem C#.

Justificativa: Por ser uma linguagem robusta e orientada a objetos. Além disso, permite o desenvolvimento rápido de aplicações, mantendo a expressividade e a elegância de uma linguagem de programação criada para o desenvolvimento de uma variedade de aplicações que executam sobre o .NET Framework.

RNF_03: O usuário não deverá precisar executar mais de 3 clicks para que todos os processamentos estejam disponíveis nas suas respectivas abas.

Justificativa: Agilizar e simplificar o processo de utilização do sistema e, conseqüentemente, a realização da perícia.

RNF_04: O usuário deverá manipular uma única interface, dividida em abas.

Justificativa: Simplificar o processo de utilização do sistema. Desta forma, o usuário terá todos os recursos do sistema visíveis em uma única tela, evitando confusões e garantindo a simplicidade de utilização do sistema.

(e) Execução

Na etapa de execução, iniciou-se o detalhamento do software proposto utilizando a definição dos algoritmos a serem implementados e da documentação do projeto (diagramas UML). Na fase de desenvolvimento, a codificação do software foi realizada. A Figura 2 apresenta as fases da execução do projeto.

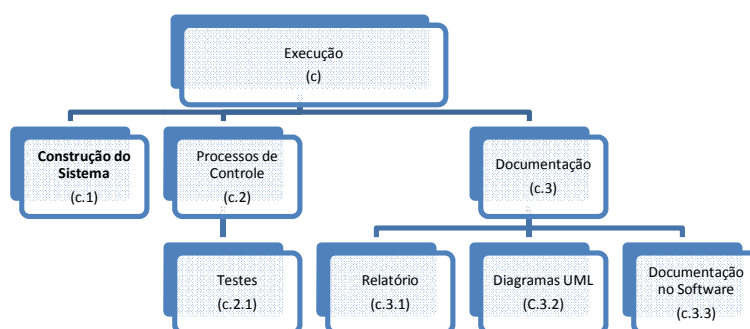


Figura 2 - Fase de execução do software *VoiceAnalysis*.

- (c.1) **Construção de Sistema *VoiceAnalysis*:** Essa etapa contemplou a codificação do software no Microsoft Visual Studio, na linguagem C#.
- (c.2) **Processos de Controle:** Para evitar problemas após a entrega da solução é necessário investir em testes, pois testes em sistemas e em documentações reduzem os riscos da ocorrência de defeitos do software (RIOS *et al.*, 2007).
 - (c.2.1) **Testes:** Essa é uma etapa de realização de diversos testes, observando a forma como são calculados os parâmetros acústicos do sinal de voz e apresentação dos resultados. E também contempla a correção dos frequentes erros no código fonte durante a implementação e a disposição dos controles da interface.
- (c.3) **Documentação:** Essa etapa descreve cada parte do código fonte. Consiste também no conjunto de manuais gerais e técnicos, além de diagramas explicando o funcionamento de um software como um todo e cada parte dele. A etapa de documentação está dividida em três etapas:
 - (c.3.1) **Relatório:** Foi confeccionado um relatório que contém os capítulos teóricos, a justificativa, os requisitos, a metodologia e os resultados do projeto;
 - (c.3.2) **Diagrama UML:** A UML propõe que a modelagem seja realizada através de vários diagramas que facilitam a visualização das características do software modelado. Para este software foram aplicados três diagramas, cada qual com sua função que em conjunto possibilitam o entendimento do funcionamento do software como um todo.

O primeiro passo da modelagem é descobrir quem são os usuários do sistema (Diagrama de Caso de Uso). Neste caso, o Perito utiliza o sistema “VoiceAnalisys” para a inserção de áudio no sistema e para visualizar os parâmetros acústicos gerados após o processamento digital do sinal da voz. A Figura 3 ilustra o diagrama de caso de uso do software.

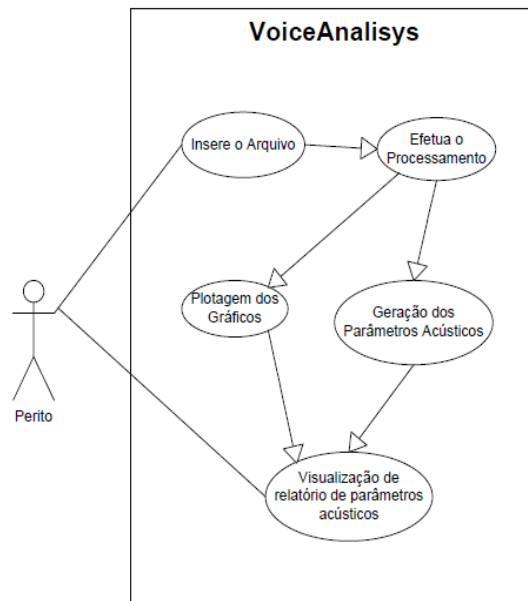


Figura 3 - Diagrama de caso de uso do software *VoiceAnalisys*.

De acordo com a Figura 3, observamos que o único usuário do sistema é o perito. Este usuário é o responsável por realizar as ações de inserir o áudio e visualizar as informações geradas após o processamento do sinal de voz.

A Figura 4 apresenta o diagrama de atividades do software. Este diagrama apresenta como a conexão entre diversas atividades são interdependentes. E também, apresenta a forma de geração dos parâmetros acústicos para comparação de duas vozes é resultado de várias operações de processamento dos sinais de voz.

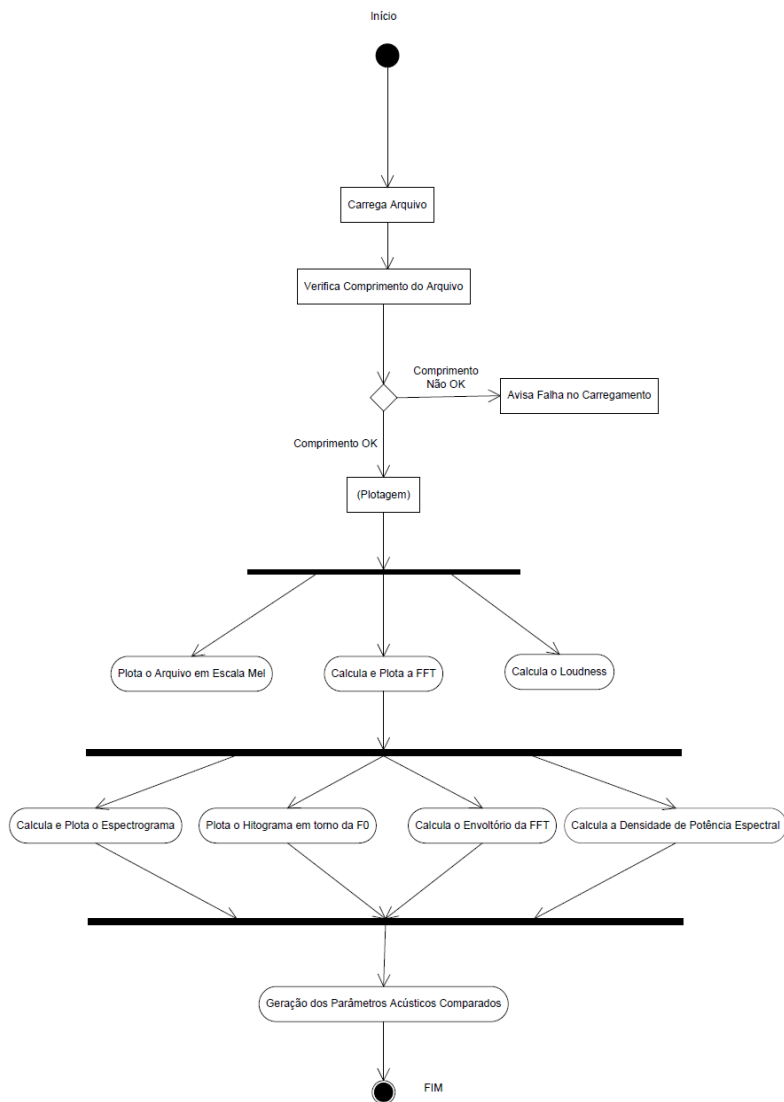


Figura 4 - Diagrama de atividade do software *VoiceAnalysis*.

(f) Encerramento

Nesta etapa são realizados testes de funcionamento, pequenas alterações/correções no software, verificação da exatidão do processamento digital dos sinais de voz e finalização do projeto.

4 Gerenciamento do projeto

O gerenciamento do projeto teve como objetivo controlar as principais restrições do projeto: escopo, tempo e qualidade. A meta durante todo o desenvolvimento foi submeter o projeto ao menor número possível de mudanças de escopo, controlar eventuais atrasos durante as etapas intermediárias e garantir que a entrega final fosse um protótipo que satisfizesse todas as especificações do projeto.

A linguagem de programação adotada foi o C#, que é uma linguagem orientada a objetos desenvolvida pela Microsoft. Sua sintaxe é baseada em C++, sendo procedural e orientada a objetos. O projeto foi implementado no ambiente de desenvolvimento Microsoft Visual Studio 2008 Professional *Edition*. Trata-se de um ambiente que provê a desenvolvedores as ferramentas e o suporte necessário à codificação na linguagem C#. Permite a criação de aplicações robustas usando toda a infra estrutura do .NET Framework. Este é um componente de software que pode ser adicionado ao sistema operacional Microsoft Windows. Ele provê um extenso conjunto de soluções previamente codificadas que cumprem os requisitos comuns para o desenvolvimento e controla a execução de programas escritos especificamente para esse framework.

Sua escolha foi feita de acordo com o grau de adequação aos requisitos/especificações do projeto e com a disponibilidade de bibliotecas já existentes para o processamento de sinais de voz. As práticas de gerenciamento adotadas neste projeto permitiram alcançar os resultados esperados, com o sistema de apoio a perícia de identificação de locutores *VoiceAnalysis* em pleno funcionamento.

5 Resultados

O resultado final do projeto foi a elaboração de um sistema de processamento de voz que permite a comparação entre parâmetros acústicos de dois sinais de voz com o intuito de apoiar a perícia de identificação de locutores: *VoiceAnalysis*. O sistema é capaz de realizar a extração de parâmetros temporais e no domínio das frequências, além da comparação estatística de parâmetros quantitativos. A Figura 5 apresenta a interface de inicialização do programa.

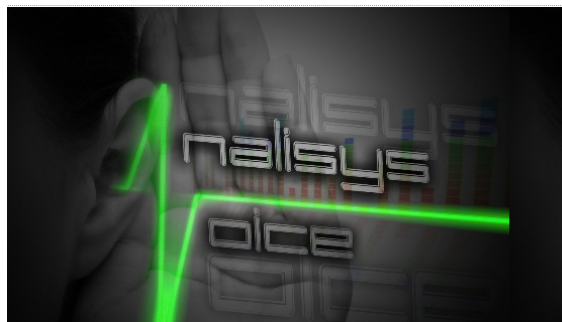


Figura 5 - Interface inicial do software *VoiceAnalysis*.

A Figura 6 mostra a interface principal do software. Nessa interface é possível executar os áudios carregados pelo software. Após a seleção dos áudios, o processamento dos mesmos já é, automaticamente, realizado.

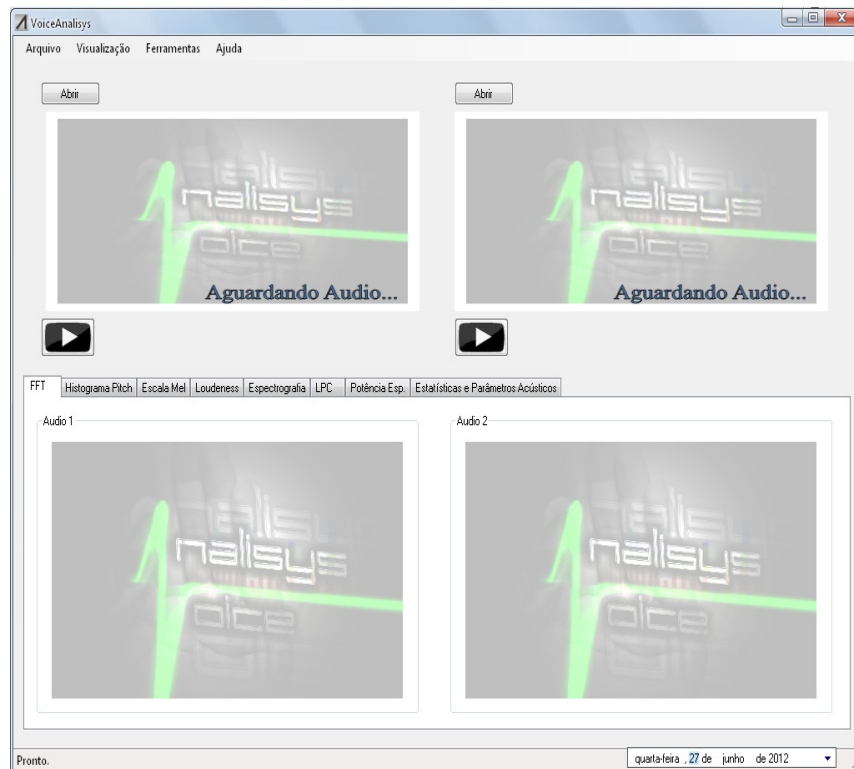


Figura 6 - Interface principal do software *VoiceAnalysis*.

A Figura 7 mostra a interface do software após a seleção dos arquivos de áudio. É possível observar que software deixa disponível na parte superior do gráfico o endereço local do arquivo, facilitando para o usuário ter certeza do áudio que está sendo processado. E também, já calcula e apresenta a FFT dos dois sinais de voz – que é a primeira aba de processamento. O limite do eixo da FFT ajustado é a metade da frequência de amostragem, que nesse e em muitos casos é de 44,1kHz, portanto os eixos da FFT variam de 0 até 22,05 kHz. No entanto, como as maiores componentes de energia estão abaixo de 5000 Hz, a ferramenta gráfica do software permite realizar zoom horizontal do gráfico da FFT e visualizar uma determinada frequência de interesse.

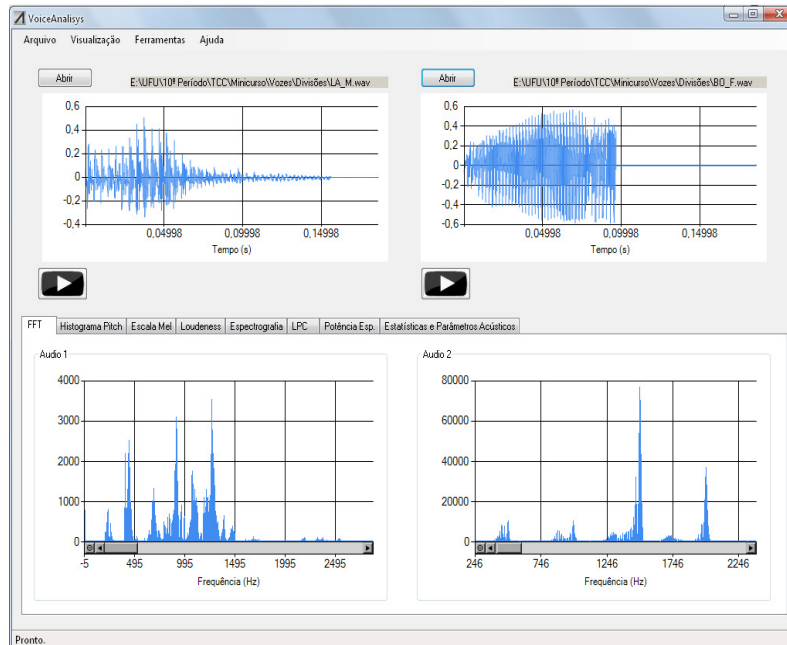


Figura 7 - Interface do Software *VoiceAnalysis* – Processamento da FFT.

A Figura 8 apresenta a visualização do Histograma de Energia em torno da Frequência Fundamental F0, em ciclos para cada componente de frequência.

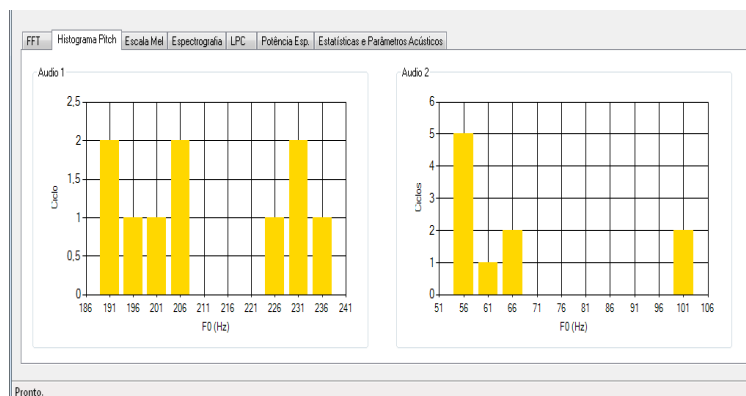


Figura 8 - Interface do Software *VoiceAnalysis* – Processamento do Histograma Pitch.

A Figura 9 apresenta a FFT convertida para a Escala Mel. Tal gráfico também pode ser aplicado o zoom.

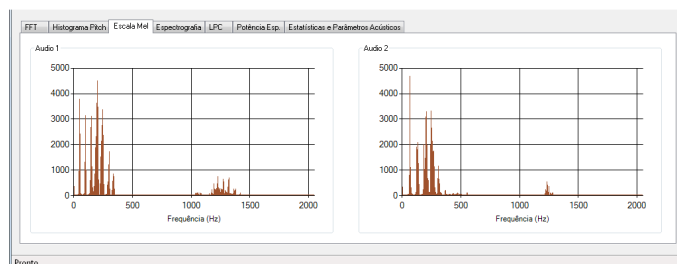


Figura 9 - Interface do software *VoiceAnalysis* – Processamento da Escala Mel.

A Figura 10 mostra a plotagem do parâmetro *loudness*, que pode ser calculado de maneira global, mas também pode ser calculado no tempo amostrando uma determinada quantidade de valores no tempo. Nesse caso, cada ponto corresponde a uma amostragem de 100 pontos do gráfico de áudio para cada ponto do gráfico de *loudness*.

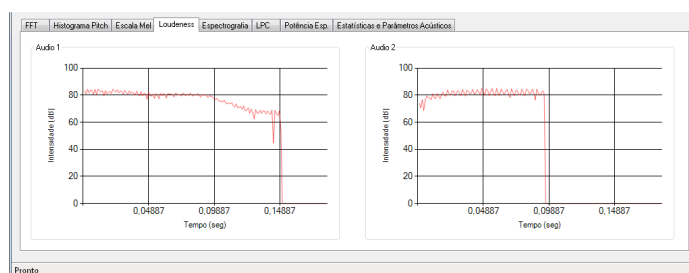


Figura 10 - Interface do software *VoiceAnalysis*– Processamento *Loudness*.

A Figura 11 mostra a representação do Espectrograma do conteúdo do áudio. Os espectrogramas são diagramas tempo-frequência que tiveram seus eixos e seus posicionamentos alinhados de forma ao usuário poder relacionar o gráfico temporal superior. O espectrograma é um gráfico que apresenta três dimensões: Frequência (eixo vertical), Tempo (eixo horizontal) e Amplitude (escala de cores). A escala de cores é apresentada da seguinte forma: quanto maior o valor em módulo da variável representada, mais escuro será o ponto correspondente, e quanto menor o módulo mais claro será a cor do ponto.

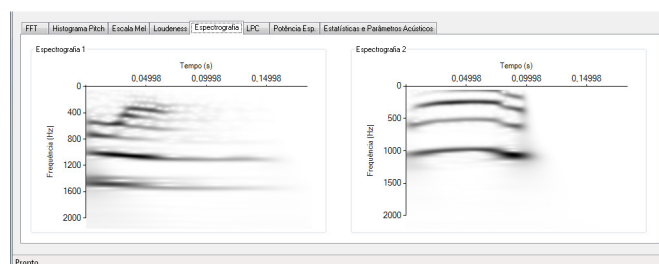


Figura 11 - Interface do Software *VoiceAnalysis*– Processamento Espectrograma.

A Figura 12 mostra o gráfico do LPC, que é, basicamente, o envoltório da FFT.

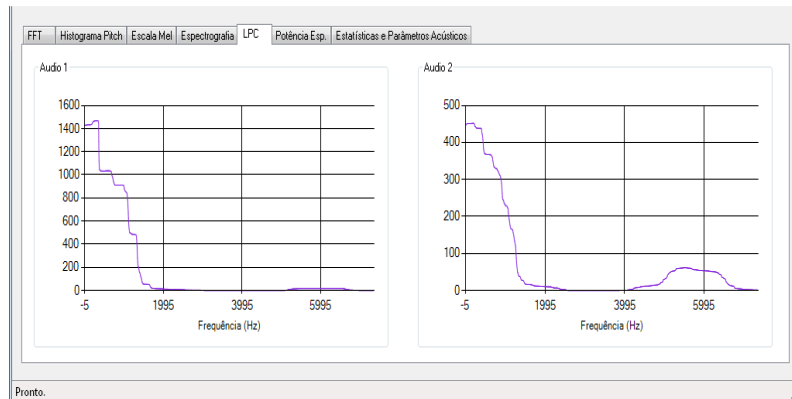


Figura 12 - Interface do software *VoiceAnalysis* – Processamento LPC.

A Figura 13 mostra o processamento referente à potência espectral, que apresenta um gráfico temporal extraído do espectrograma. Esse gráfico mostra a energia espectral ao longo do tempo.

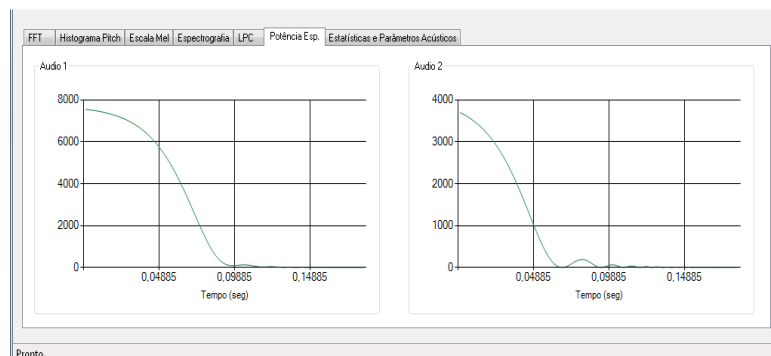


Figura 13 - Interface do software *VoiceAnalysis*– Processamento Potência Espectral.

A Figura 14 apresenta as estatísticas e parâmetros acústicos. Esta interface extrai parâmetros quantitativos dos gráficos anteriores para os dois áudios e realiza uma comparação na coluna “erro”. Essa coluna é colorida de acordo com parâmetros de erros pré-estabelecidos, nos quais valores de erro menores que 10% são verdes, os menores que 30% são laranja e os maiores que 30% são vermelhos. Esta representação tem como objetivo fornecer uma análise rápida, uma vez que a decisão do quanto os valores de erro entre os áudios poderão variar para influenciar o resultado cabe ao perito. Essa interface também permite exportar esses parâmetros quantitativos para um relatório no Microsoft Word, facilitando, assim, a confecção do laudo pericial (Figura 15).

Parâmetros	Audio1	Audio2	Erro
Pitch - F0 (Hz)	215	295	15,63 %
Formante - F1 (Hz)	410	415	1,20 %
Formante - F2 (Hz)	1175	1390	15,47 %
Formante - F3 (Hz)	2000	1860	7,93 %
Loudness (dB)	98,35	102,78	4,31 %
Shimmer	17,4	8,53	103,99 %
Jitter	5	215	97,67 %
Vozeamento (seg)	0,02834	0,08730	67,54 %

Parâmetros	Audio1	Audio2	Erro
Mediana F0	220	40	490 %
Desvio Padrão DEP	15708,95	38333,07	59,02 %
Desvio Padrão LPC	5288,21	45769,31	88,45 %
Média DEP	485,11	1583,76	69,37 %
Média LPC	12,77	110,33	88,43 %
-			
-			
-			

Figura 14 - Interface do software *VoiceAnalisys*– Processamento das Estatísticas e Parâmetros Acústicos do sinais de áudio.

VOICEANALISYS- RELATÓRIOS
domingo, 24 de junho de 2012

PARÂMETROS ACÚSTICOS

Audio1:

Pitch - F0 (Hz): 131
 Formante - F1 (Hz): 371,45
 Formante - F2 (Hz): 454,89
 Formante - F3 (Hz): 1905,69
 Loudness (dB): 100,72
 Shimmer: 16,1
 Jitter: 0
 Vozeamento (seg): 0,06236
 Desvio Padrão DEP: 98
 Desvio Padrão LPC: 759,90
 Média DEP: 4557,44
 Média LPC: 99,80
 Mediana F0: 3,06

Audio2:

Pitch - F0 (Hz): 107
 Formante - F1 (Hz): 368,76
 Formante - F2 (Hz): 686,37
 Formante - F3 (Hz): 1816,86
 Loudness (dB): 99,88
 Shimmer: 16,25
 Jitter: 2
 Vozeamento (seg): 0,11905
 Desvio Padrão DEP: 82
 Desvio Padrão LPC: 9607,51
 Média DEP: 3101,42
 Média LPC: 503,64
 Mediana F0: 1,97

Página: 1 de 1 Palavras: 108 Português (Brasil) 80%

Figura 15 - Relatório com os parâmetros acústicos calculados pelo *VoiceAnalisys*, exportado para o Microsoft Word

6 Discussão

A interface principal do software desenvolvido possui todos os componentes escritos na Língua Portuguesa e um esquema de abas que facilita a navegação pelo sistema. O software apresenta uma quantidade mínima de botões, também com o intuito de simplificar a utilização.

Nos processamentos dos sinais, na aba FFT, a mesma fornece o gráfico dos dois áudios no momento em que o áudio é carregado e permite o usuário realizar zoom no gráfico se assim desejar, tal funcionalidade disponível em praticamente todos os outros gráficos desse sistema. Os gráficos derivados da FFT, histograma F0, escala mel e LPC, também são mostrados de forma comparada com o objetivo de respaldar no laudo os valores quantitativos apresentados. Quanto ao espectrograma, o mesmo é apresentado da maneira mais usual possível, tempo na horizontal e frequências na vertical. Esta visualização permite ao perito relacionar com o gráfico da voz no tempo. E ainda, os gráficos no domínio do tempo, *loudness* e potência espectral, alinhados com o gráfico de voz respaldam o laudo pericial e ainda permitem análises entre os gráficos e acompanhar as variações ao longo do tempo.

Resumindo todos esses gráficos de processamentos, a aba Estatísticas e Parâmetros Acústicos tornam os principais parâmetros quantitativos dos gráficos disponíveis ao perito. Tal funcionalidade, não é disponibilizada nas atuais soluções periciais, tendo o perito que extrair relatórios em softwares separados ou até mesmo extrair esses valores apenas com a visualização do gráfico gerado pelo software utilizado na perícia. Assim, essa funcionalidade torna o *VoiceAnalysis* uma ferramenta útil e para as perícias de voz.

Além da extração quantitativa dos parâmetros acústicos, o *VoiceAnalysis* provê a visualização desses parâmetros no formato de tabelas, acrescentado a essas tabelas uma coluna que efetua a comparação dos valores dessas tabelas. Desta forma o software extrai os parâmetros quantitativos de duas amostras de voz nas quais se deseja comparar, mostra os gráficos e ainda compara os parâmetros dos gráficos. Complementando a visualização desses fatores, o *VoiceAnalysis* provê uma análise do erro entre os parâmetros. Funcionalidades estas que apresentam ao usuário cores diferentes, sinalizando a proximidade ou não dos valores comparados.

Quanto à utilização do *VoiceAnalysis* em perícias pode-se destacar a facilidade e a rapidez na extração dos parâmetros pelo perito para análise. Isso porque, após carregar os gráficos, o perito já possui todos os parâmetros calculados, automaticamente, na aba Estatísticas e Parâmetros Acústicos e isenta o perito de ter que extrair essas informações, facilitando e agilizando o trabalho pericial.

Portanto, a análise das aplicações do software *VoiceAnalysis* permite inferir o alcance dos objetivos propostos, que é a do desenvolvimento de uma ferramenta computacional de apoio a perícia de voz. Tal ferramenta supre algumas demandas da perícia de voz, como a comparação gráfica e principalmente a extração de parâmetros acústicos e comparação entre os mesmos em tabelas.

Assim, a proposta do *VoiceAnalysis* surge como um diferencial dentre os *softwares* atualmente utilizados nos trabalho de perícia de confronto de voz. Nenhum dos sistemas computacionais existentes contempla todas as facilidades que o *VoiceAnalysis* oferece para o perito.

Conclusões

A fala pode ser utilizada na identificação de pessoas, pois a mesma é um dado biométrico e, portanto, é uma característica biológica que pode ser determinada quantitativamente. A fonética, que é uma ciência que estuda a física relativa aos sons da fala humana, também é utilizada para fins jurídicos. Quanto à metodologia atual dos exames de locutores numa perícia de identificação forense da voz, as mesmas se baseiam numa análise perceptual e acústica.

Atualmente, o padrão-ouro utilizado para avaliação de voz é a avaliação perceptivo-auditiva, que é subjetiva e depende da experiência do avaliador (BELE, 2005). Por isso, diante da subjetividade da avaliação perceptivo-auditiva, cada vez mais, têm-se utilizado ferramentas objetivas complementares para avaliação da voz, como a análise acústica. Porém, os poucos softwares disponíveis no mercado para apoiar a perícia de confronto de voz utilizam não contemplam todos os parâmetros de análise acústica necessários neste tipo de trabalho e não possuem uma interface intuitiva e com rápido acesso/processamento por parte do usuário, o que torna o trabalho do perito muito dispendioso.

Nesse contexto, neste projeto foi proposto um *software* de apoio à perícia de voz, com uma interface mais amigável ao perito, e que contenha os principais parâmetros utilizados no trabalho pericial, facilitando assim as comparações estatísticas entre dois áudios. Assim sendo, o sistema proposto neste trabalho realiza uma análise acústica de dois sinais de voz que servirá de apoio ao perito, pois realiza em uma única ferramenta todo o processamento necessário para a perícia de voz. Além disso, o sistema implementa a unificação da análise quantitativa comparada, trabalho que o perito precisa realizar e que, até então, não existia nenhuma ferramenta computacional que tornasse essa análise e extração mais ágil durante uma perícia. Tais funcionalidades no *software* foram implementadas no intuito de otimizar o processo pericial de análise acústica, o qual demanda muito tempo do perito para analisar em diversas ferramentas e depois sintetizar todas as informações pertinentes no seu laudo.

Os trabalhos futuros que podem ser realizados a partir dessa proposta são:

- Realização de análises em tempo real;
- Inserção de mais parâmetros acústicos e estatísticos para comparação;
- Eliminação de ruídos do sinal de voz;
- Seleção dos trechos a serem analisados;
- Armazenamento das tabelas em banco de dados para análise futura;
- Exportação das tabelas para o Microsoft Excel;
- Incorporação da possibilidade do perito incluir pareceres perceptuais e transcrição fonética;
- Tomada de decisão automática, com base na análise acústica, informando se os dois sinais de voz foram ou não produzidos pelo mesmo aparelho fonador.

SILVA, G. M.; SÁ, A. A. R.; FARIA, V. N. R.; SOARES, A. B.; DIB NETO, I. K. G.; PEIXOTO, H. Computer interface of biometrics as a tool to support the expertise of clash of voice. *Rev. Bras. Biom.*, São Paulo, v.30, n.4, p.442-460, 2012.

■ **ABSTRACT:** *The Biomedical Engineering Forensic uses the assignments of Biomedical Engineering for legal purposes, encompassing issues addressed by the biometrics human identification. As biometrics a science that promotes processes of identifying individuals based on physiological or behavioral characteristics, presenting several applications, among them the elucidation of crimes. Among the various types of existing biometrics, biometrics voice consists of an identification process that assesses the individual's speech, as well as their behavioral characteristics during speech. In this context, Forensic Phonetics is a science that seeks to determine the authenticity of audio files presented to investors of the Law. This subject doesn't present their current methodologies and techniques completely grounded. This work presents an investigation of the entire history of this science, checks the current point of development in the same place and a review of current speaker recognition techniques. In addition, this paper proposes the development of a computational tool that integrates multiple complementary acoustic parameters of speech, a simple interface to expedite recognition of the expertise of the speaker. The results showed that the tool is efficient representation of a comparison between the acoustic parameters of the analyzed audio.*

■ **KEYWORDS:** *Biometrics; voice identification; forensics; forensic engineering.*

Referências

BEHLAU, I. R. M. *Percepção da fala: análise acústica*. São Paulo, 1993. 57p.

BEIGI, H. Speaker recognition. Intech Open Access Publisher: *Biometrics Book*. 2011, p.3-28.

BOERSMA, P.; WEENINK, D. Praat: version 5.3.12. Amsterdam. 2012. Disponível em: <<http://www.fon.hum.uva.nl/praat>> Acesso: 7 Abr. 2012.

BORBA, F. S. *Introdução aos estudos linguísticos*. São Paulo: Companhia Editora Nacional, 1975. 251p.

BRAID, A. C. M. *Fonética forense: identificação de falante*. *Rev. Prova Material*. Salvador, 2004.

BRAID, A. C. M. *Fonética forense*. Porto Alegre: Sagra Luzzatto, 1999. 129p.

FIGUEIREDO, R. M. *Identificação de falantes: aspectos teóricos e metodológicos*. 1994. Tese (Doutorado) – Universidade Estadual de Campinas, Campinas, 1994.

FLORES, J. F. V. C. *Novas contribuições à verificação automática de locutor para fins forenses*. 2008. Dissertação (Mestrado) – Instituto Militar de Engenharia, Rio de Janeiro, 2008.

FONOVIEW, Software Fonoview Versão 4. Pato Branco - Paraná. Disponível em: <www.ctsinformatica.com.br>. Acesso em: 20 Mar. 2012.

MASON, J. S.; BRAND, J. D. *The role of dynamics in visual speech biometrics*. *Proc. Speech, Lang. Law*, v.12, n.2, 2005.

MORISSON, A. L. C. *Identificação humana pela voz*. (2004) Disponível em: <<http://www.apcesp.com.br/idvoice.htm>>. Acesso em: 5 Mar. 2012.

- RABINER, L. R.;JUANG, B. H. *Fundamental of speech recognition*. New Jersey: Prentice-Hall International, 1993. 507p.
- RABINER, L. R.; SCHAFER, R. W. *Digital processing of speech signals*. Englewood Cliffs: Prentice-Hall, 1978. 528p.
- RIOS, E.; CRISTALLI, R.; MOREIRA, T.; BASTOS, A. *Base de conhecimento em teste de software*. 2.ed. São Paulo: Martins Fontes, 2007.
- SCATENA, H. J. *A física aplica à perícia criminal: fonética forense*. 2010. 32f. Monografia (Trabalho de Graduação em Física) – Universidade Católica de Brasília, Brasília, 2010.
- TOSI, O. *Voice identification: theory and legal applications*. Baltimore: University Park Press, 1979.
- VELHO, A. J; GEISER, C. G.; ESPINDULA, A. *Ciências forenses:uma introdução às principais áreas da criminalística moderna*. Campinas: Milenium, 2012. 416p.
- VOXMETRIA. Software Voxmetria Versão 4. Pato Branco, Paraná. Disponível em: <www.ctsinformatica.com.br>. Acesso em: 18 Mar. 2012.
- WOLF, J. Efficient acoustic parameters for speaker recognition. *J. Acoustical Soc. Am.*, Melville, v.51, n.6B, p.2044, 1972.

Recebido em 13.08.2012

Aprovado após revisão em 22.04.2013