

Contribuições para o Enhanced Digital Radio Profile (EDR) do Ginga destinado para uso no Sistema Brasileiro de Rádio Digital. Essas contribuições serão parte integral da dissertação de mestrado do autor.

A motivação desse trabalho é colaborar para a definição de um perfil do Ginga para o SBRD que contenha um suporte melhorado para o tratamento de mídias aurais e o suporte à recursos afinados à característica móvel de recepção do rádio.

Novos recursos permitirão um aumento no valor agregado do conteúdo radiofônico, novas possibilidades para anúncios e serviços além de serviços de áudio para pessoas com restrições de visão e analfabetos (33 milhões de brasileiros são analfabetos funcionais). Também um suporte melhorado à execução de aplicações Ginga em receptores sem tela e considerando sempre o contexto de banda passante limitada do rádio.

- Novas mídias: SSML e SVG
- Suporte a áudio 3D
- Novos efeitos de transição
- Introdução à efeitos de áudio
- Suporte à geolocalização via coordenadas de latitude e longitude
- Suporte dependente de sistema (DRM, HD Radio)
- Utilização da linguagem NCL definida na H.761

Novas mídias: SSML e SVG. Speech Synthesis Markup Language

Suporte ao tipo de mídia XML Speech Synthesis Markup Language (SSML), que é recomendação W3C para sintetização de voz, já suportado por vários renderizadores TTS livres como o Festival e eSpeak.

SSML

- Extensão: .ssml
- Tipo Mime: “application/ssml+xml”
- Versão atual: 1.0/2004 (1.1/2010 Candidate Recommendation)

Novas mídias: SSML e SVG. Speech Synthesis Markup Language

O SSML é parte do VoiceXML, que é o padrão W3C para Voice browsers, conjuntamente aos seguintes padrões:

- Speech Recognition Grammar Specification (SRGS)
- Semantic Interpretation for Speech Recognition (SISR)
- Pronunciation Lexicon Specification (PLS)
- State Chart XML (SCXML)

Uma proposta seria integrar não somente o SSML mas todo o VoiceXML, de forma a adicionar uma nova mídia de apresentação, de forma semelhante à HTML e SMIL, ao Ginga. Existe uma implementação aberta de browser VoiceXML chamada Vocalocity's OpenVX da Carnegie Mellon University.

O VoiceXML permite um controle total da aplicação embarcada via som, permitindo o acesso à pessoas com restrições visuais ou de movimento. Seria o primeiro middleware utilizado em radiodifusão a integrar tal recurso.

Novas mídias: SSML e SVG. Speech Synthesis Markup Language

Propriedades do SSML

- **soundLevel**, **trebleLevel**, **bassLevel**, **freeze**
- **language**: Define ou sobrescreve a língua definida no documento SSML elemento raiz <speech>. Por exemplo para ensino de línguas. Outras opções são “system” para usar a língua definida em *system.language* ou “”.
- **voiceGender**: “male”, “female”, “neutral” or “”.
- **voiceAge**: Um inteiro não negativo.
- **voiceVariant**: Um inteiro positivo. Por exemplo “2” indica a segunda voz de um homem ou mulher, por exemplo.
- **voiceName**: Voz específica. Uma lista separada por espaço com as vozes preferidas. “liane”, “<person's name1>... <person's name x>” or “”.

Novas mídias: SSML e SVG. Speech Synthesis Markup Language

Atributos redefinidos para área de mídias SSML

- **beginText**: ID do elemento SSML. Refere-se ao momento de início da sintetização do conteúdo embutido no elemento com dado ID.
- **endText**: ID do elemento SSML. Refere-se ao momento do final da sintetização do conteúdo embutido no elemento com dado ID.

Suporte ao tipo de mídia XML Scalable Vector Graphics (SVG), perfil “Basic”, mandatório, ou “Tiny”, recomendado. SVG é o formato recomendado pela W3C para imagem vetorial. Já é amplamente suportado por softwares livres como o Cairo.

SVG

- Extensão: .svg ou .svgz
- Tipo Mime: “image/svg+xml”
- Versão atual: 1.1/2011

Propriedades do SVG

- As mesmas propriedades existentes para elementos visuais serão utilizadas. A modificação de propriedades que aumentam ou diminuem as dimensões visuais da mídia serão feitas sem perda de fidelidade. Caso suportadas, animações SVG terão o mesmo comportamento da mídia de vídeo.

O perfil “Basic” é mandatório e não inclui animações e o perfil “Tiny”, que inclui animações, é recomendado.

Novas propriedades para a mídia de áudio que permitirão a montagem de cenais aurais 3D que deverão ser renderizadas aplicando head-related transfer function (Batteau, 1967) para ambos os ouvidos (H_l e H_r):

$$H = f(\omega, \theta, \phi, d)$$

Na qual ω é a frequência angular da fonte sonora, θ é o azimute, ϕ é a elevação e d é a distância do centro da cabeça do ouvinte à fonte de som.

No caso de fones de ouvido, a saída da HRTF é reproduzida por cada falante do fone, no caso de caixas de som discretas, é necessário distribuir os dois canais entre as caixas de som do ambiente de forma a garantir que os dois ouvidos do ouvinte recebam os sinais de áudio de forma a manter os efeitos do filtro aplicado, devendo-se introduzir tratamento no áudio de forma a atingir tal objetivo.

Novas propriedades monoaurais

- **azimuth**: Azimute. Valores possíveis: [angle] ou left-side, far-left, left, center-left, center, center-right, right, far-right, right-side, com ou sem behind. [angle] variando de “-360deg” até “360deg”. Exemplos: “behind far-right”, “90deg” (igual a “right-side”).
- **elevation**: Elevação. Valores possíveis: [angle] ou below, level, above, com [angle] variando de “-90deg” até “90deg”.
- **distance**: Distância da fonte de áudio até o ouvinte. Valores não negativos, em metros. Exemplo: “10m”.

No caso de uma mídia com dois ou mais de dois canais, caso alguma propriedade de espacialidade seja utilizada, todos os canais serão somados e tratados como uma mídia monoaural.

A renderização utilizando HRTF é uma recomendação, podendo ser utilizados outros métodos que provejam resultados subjetivos melhores como Wave Field Synthesis (Berkhout, 1988).

Novos efeitos de transição

Na mesma linha das transições de vídeo do NCL, existe a transição type “fade” e subtype “fade”. Será adicionada a transição type “slowDown” e “speedUp”, de subtypes “soft” ou “hard”, que são semelhantes ao efeito de acionamento ou desligamento do motor do toca discos, a partir do disco parado ou girando na velocidade de nominal de operação (33, 45 ou 72rpm). No subtype “soft” a aceleração ou desaceleração da reprodução é mais suave, enquanto no “hard” é mais abrupta.

Atributos de transição de áudio

- **id, type, subtype, dur, startProgress, endProgress**

No caso das transições de “speedUp” e “slowDown” o tempo de duração da transição irá determinar o quão rápido será o aumento da velocidade ou a redução da mesma.

Introdução à efeitos de áudio

Serão introduzidos efeitos de áudio que poderão ser aplicados tanto a algum fluxo principal do áudio da emissora (max. 4 no DRM e 3 no HD Radio) quanto em mídias transmitidas via carrossel ou baixadas via URI da rede de canal de retorno. Será definida uma nova propriedade para as mídias de áudio de nome `audioEffect`. Os valores possíveis são do tipo lista no seguinte formato (type, param1, param2, param3, ...):

Valores do `audioEffect`

- **echo**, delay, feedback, intensity. Echo or (simple) reverb effect to an audio stream. Delay of the echo in nanoseconds. Allowed values: ≥ 1 . Default value: 1. Amount of feedback. Allowed values: $[0,1]$. Default value: 0. Intensity of the echo. Allowed values: $[0,1]$. Default value: 0. `gst-audiofx`.

Valores do audioEffect

- **compansion**, mode, characteristics, ratio, threshold. Act as a compressor or expander. Mode selects between compressor or expander. Ratio that should be applied. Allowed values: ≥ 0 (float). Default value: 1. Threshold until the filter is activated. Allowed values: $[0,1]$. Default value: 0.
gst-audiofx.
- **highPass**, cutoff, transitionBand. Highpass filter. Cutoff frequency, in Hz. Transitionband width, in Hz.
- **lowPass**, cutoff, transitionBand. Lowpass filter. Cutoff frequency, in Hz. Transitionband width, in Hz.
- **bandPass**, lowerFreq, upperFreq, transitionBand. Band-pass filter. Lower frequency in Hz. Upper frequency in Hz. Transitionband width, in Hz.

Valores do audioEffect

- **bandReject**, lowerFreq, upperFreq, transitionBand.
Band-Reject filter. Lower frequency in Hz. Upper frequency in Hz. Transitionband width, in Hz.
- **panorama**, pan, method. Stereo panorama effect with controllable pan position. Position in stereo panorama (-1.0 left ->1.0 right). Allowed values: [-1,1]. Default value: 0. Panning method: “psychoacoustic” or “simple”. gst-audiofx.

Valores do audioEffect

- **flanger**, frequency, lrPhaseShift, depth, delay, contour, dryLevel, wetLevel. Flanger and Chorus effect. This plugin is an implementation capable of creating traditional Chorus and Flanger effects, spiced up a bit to make use of stereo processing. It sounds best on guitar and synth tracks.
<http://tap-plugins.sourceforge.net/ladspa/chorusflanger.html>.
- **reverb**, Decay [ms], Dry Level [dB], Wet Level [dB], Comb Filters [ON, OFF], Allpass Filters [ON, OFF], Bandpass Filter [ON, OFF], Enhanced Stereo [ON, OFF], Reverb Type [0, 42].
<http://tap-plugins.sourceforge.net/ladspa/reverb.html>;

Suporte à geolocalização via coordenadas de latitude e longitude

Atualmente o Ginga suporta geolocalização utilizando o CEP, informação disponível no nó settings. Faz-se necessário uma informação sobre a geolocalização mais detalhada para as aplicações no contexto do rádio, majoritariamente executadas num contexto de mobilidade.

Ao menos duas possibilidades de inserir o suporte no NCL. Uma seria via nó settings, utilizando as propriedades.

Propriedades do nó settings

- **latitude**
- **longitude**
- **precision**

Suporte à geolocalização via coordenadas de latitude e longitude

Outra possibilidade é adicionar uma mídia especial, de tipo:

Tipo MIME da mídia de geolocalização

application/x-ncl-geolocalization

Quando iniciada, o sistema de geolocalização do aparelho pode ser ligado, sendo sinalizado via nó settings propriedade system.hasGPS sua presença.

Propriedades do mídia de geolocalização

- **latitude**
- **longitude**
- **precision**

Suporte à geolocalização via coordenadas de latitude e longitude

Além disso será possível definir áreas no elemento de mídia de geolocalização que serão utilizadas em conectores no papel “onBegin”, caso o ouvinte entre na região, ou “onEnd” caso o ouvinte saia da região.

O Ginga se tornará o primeiro middleware que permite o uso de informações de geolocalização em ambiente declarativo.

Suporte dependente de sistema (DRM, HD Radio)

Devido aos diferentes conjuntos de recursos dos dois padrões de rádio digital sendo considerados para adoção no Brasil, algumas diferenças em como localizar objetos no multiplex via URI e quais informações estarão presentes no nó settings serão diferentes. Também a forma de multiplexar os arquivos da aplicação Ginga irão diferir.

A parte de planos gráficos atualmente no Ginga não será utilizada, ficando somente um plano default disponível.

Novos codecs sem royalties deverão ser suportados, como Opus para áudio, VP9 para vídeo e WebP para imagens.

Definir integração com Journaline, definir identificação única do aplicativo, suporte a envio de comandos de edição e acesso aos elementos presentes no broadcast.

Utilização da linguagem NCL definida na H.761

Ao utilizar o NCL e o Ginga-NCL definido na H.761 como base do novo perfil EDR, uma versão mais nova do NCL será utilizada (3.1) e novos recursos NCLua estarão disponíveis, além do novo perfil Raw também poder ser utilizado como opção ao EDTV. Devido ao crescimento geométrico da capacidade de processamento dos dispositivos móveis não existe a necessidade de uma otimização em termos da definição da linguagem.