



UNIVERSIDADE FEDERAL DA PARAÍBA
CENTRO DE CIÊNCIAS APLICADAS E EDUCAÇÃO
DEPARTAMENTO DE CIÊNCIAS EXATAS
BACHARELADO EM SISTEMAS DE INFORMAÇÃO



CineLIBRAS: Uma Proposta para Geração Automática e Distribuição de Janelas de LIBRAS em Salas de Cinema

Leonardo de Araújo Domingues

Rio Tinto - PB
Outubro de 2013

Leonardo de Araújo Domingues

**CineLIBRAS: Uma Proposta para Geração
Automática e Distribuição de Janelas de LIBRAS em
Salas de Cinema**

Trabalho de Conclusão de Curso
apresentado ao Centro de Ciências
Aplicada e Educação da
Universidade Federal da Paraíba,
para a obtenção do título de
Bacharel em Sistemas de
Informação.

Orientador: Prof. Dr. Tiago Maritan Ugulino de Araújo

Rio Tinto - PB

Outubro de 2013

D671c Domingues, Leonardo de Araújo.

CineLIBRAS: Uma Proposta para Geração Automática e Distribuição de Janelas de LIBRAS em Salas de Cinema / Leonardo de Araújo Domingues. – Rio Tinto: [s.n.], 2013.

83f.: il. –

Orientador: Tiago Maritan Ugolino de Araújo.

Monografia (Graduação) – UFPB/CCAIE.

1. Tecnologia da Computação. 2. Acessibilidade. 3. Língua Brasileira de Sinais. 4. Cinema Digital. 5. CineLIBRAS – Desenvolvimento. I. Título.

UFPB/BS-CCAIE

CDU: 004 (043.2)

*À minha família, que sempre me incentivou a
nunca, em hipótese alguma, desistir dos
meus sonhos.*

Agradecimentos

Primeiramente, agradeço a Deus, por me conceder o dom da vida e a capacidade de desenvolver meus talentos, assim também como a saúde e a força necessária nos momentos de desalento dessa jornada.

De modo muito especial, agradeço a minha mãe Elizabeth, pelos seus ensinamentos tão valiosos, pelo incentivo nos momentos em que estive desanimado, por toda compreensão quando mais precisei, e por me ensinar desde muito cedo que o caráter do homem revela a sua integridade. Enfim, agradeço por todo seu amor, carinho e proteção.

À minha avó Terezinha, pela sua serenidade e bondade de espírito, fé em Deus e confiança na recompensa após o trabalho honesto. E ainda, pelas nossas conversas tão alegres e felizes nos fins de tarde.

Aos meus amados irmãos, Rafael e Beatriz, pelo total apoio e encorajamento nessa minha caminhada. Saibam que sou muito grato pelo estímulo que vocês me deram, pois, mesmo sendo imperceptível, em muitos momentos difíceis, encontrei em vocês uma razão para persistir com meus objetivos.

Ao meu grande amigo Antônio Gonçalves, que quando mais precisei, pude contar com seu auxílio e suas palavras de conforto e ânimo. Ao meu tio Carlos Henrique (*in memoriam*), que mesmo quando eu não passava de um moleque que sonhava em ser um jogador de futebol, ele já me ensinava a importância de conquistar uma formação acadêmica e seguir uma carreira para ser bem sucedido. Ao meu eterno amigo Josué (*in memoriam*), que me acolheu na condição de seu filho e me ensinou que um homem de caráter deve ser sempre fiel aos seus valores.

À minha querida Jéssyca, por sua surpreendente compreensão nos momentos em que precisei estar ausente, pelo seu amor tão confortador nas horas de angústia, e pelo carinho completamente original, que sem dúvida alguma, teve um papel fundamental para a conquista desse trabalho.

Ao meu orientador e professor Dr. Tiago Maritan, que incessantemente se fez presente quando tive alguma dúvida em relação ao desenvolvimento deste trabalho, além de outras questões acadêmicas. Pelo seu exemplo profissional tão evidente, pois, ser um educador não significa simplesmente ensinar algum conteúdo literário, mas formar pessoas de maneira que possam fazer a diferença. Enfim, agradeço por todas as cobranças e críticas, pois essas com certeza me contribuíram para a minha formação.

Aos professores do Departamento de Ciências Exatas do Campus IV, de modo especial aos professores Gilberto Farias, Carlos Hacks, Yuri Malheiros, Jorge Gabriel, Rafael Magalhães, Hermann, Raoni, Rodrigo Vilar, Juliana Aragão, Rodrigo Rebouças e Alisson Brito. E ainda, de forma muito especial, agradecer a professora e amiga Ayla Rebouças, que desde os momentos iniciais da minha formação, sempre acreditou no meu potencial. Muito obrigado pelo seu entusiasmo e determinação, eles sempre foram uma referência pra mim.

Aos meus companheiros, não menos que guerreiros, do Laboratório de Aplicações de Vídeo Digital (LAViD), Guido Lemos, Felipe Lacet, Danilo Assis, Felipe Lemos, Gutenberg, Vandhuy, Lucenildo, Fernando Brito, Luis Henrique, Igor Amorim, Hozana, Yúrika, Eduardo Lucena, Eduardo Victor, Leonardo Dantas, Virgínia, Eliza, Manuela, Manoel Gomes, Erick, Kellyane, Cheylla, Derzu, Alexandre Nóbrega, dentre outros colegas, pelos momentos de fraternidade e pelas valiosas ideias que contribuíram de alguma forma para a realização deste trabalho.

Por fim, agradeço a todos os meus familiares, amigos e colegas que, mesmo estando distante ou até mesmo que não tenham sido citados aqui explicitamente, sempre torceram por mim e me deram grande força para continuar nessa jornada. Que Deus abençoe a todos.

Resumo

Os surdos se deparam com muitas dificuldades para participar da vida em sociedade, para se comunicar com outras pessoas e acessar informações. Sua participação em alguns ambientes sociais como cinemas, teatros, ambientes educacionais, dentre outros, muitas vezes é limitada devido a ausência de soluções acessíveis que viabilizem sua participação e interação nesses ambientes. Quando disponíveis, os recursos de acessibilidade para este público se restringem à utilização de legendas no formato de texto. No entanto, os surdos possuem muitas dificuldades para ler e escrever na língua oral do seu país, uma vez que essas línguas são baseadas em sons. Portanto, para tentar reduzir esses problemas, este trabalho tem como objetivo o desenvolvimento de uma solução computacional denominada CineLIBRAS. A proposta é que janelas de Língua Brasileira de Sinais (LIBRAS), representadas por um avatar 3D, sejam geradas automaticamente a partir das legendas dos conteúdos e distribuídas dentro das salas de cinema. Dessa forma, os surdos brasileiros poderão acompanhar a tradução do conteúdo na sua língua natural de comunicação, a LIBRAS, por meio de um dispositivo móvel capaz de receber fluxos de vídeo.

Palavras-chave: Acessibilidade, Cinema Digital, Língua Brasileira de Sinais, CineLIBRAS.

Abstract

The deaf people face many difficulties to participate in the social life, to communicate with other people and to access information. Their participation in some environments as cinemas, theaters, educational environments, among others, is often limited due to the lack of solutions that enable their interaction and participation on these environments. When available, the accessibility resources for this audience is restricted to the use of subtitles in text format. However, the deaf have many difficulties in reading and writing the oral language of their country, since these languages are based on sounds. Therefore, to try reduce these problems, this work aims to develop a computational solution called CineLIBRAS. The proposal is that windows in Brazilian Sign Language (BSL), represented by a 3D avatar, are automatically generated from the contents of subtitles and distributed inside the cinema. Thus, the Brazilian deaf will follow the translated content in their natural language, the BSL, using a device mobile to receive video streams.

Keywords: *Accessibility, Digital Cinema, Brazilian Sign Language, CineLIBRAS.*

Lista de Figuras

Figura 1. Deficiência auditiva: dados da população mundial (WHO 2013)	17
Figura 2. Alfabeto na Língua Brasileira de Sinais	31
Figura 3. Arquitetura proposta por Tambassia et al. (2012)	37
Figura 4. Visão esquemática da solução da SONY (2012)	39
Figura 5. Exibição das legendas durante um jogo de Futebol Americano no <i>Dallas Cowboys Stadium</i>	41
Figura 6. Modelo conceitual da infraestrutura do CineLIBRAS	44
Figura 7. Sequência de execução para a geração e distribuição dos vídeos acessíveis .	45
Figura 8. Estrutura da mensagem utilizada na comunicação	48
Figura 9. Sequência de mensagens trocadas entre o CineLIBRAS e o Fogo Player	50
Figura 10. Arquitetura do CineLIBRAS	52
Figura 11. Estrutura do arquivo de legendas recomendado pela DCI	53
Figura 12. Modelo conceitual do experimento realizado	60
Figura 13. Grau de escolaridade dos vinte usuários surdos que participaram da primeira etapa do experimento	61
Figura 14. Momentos durante a realização da primeira etapa do experimento: (a) Alguns usuários que participaram do teste, (b) Fase de preparação dos usuários, (c) Durante a avaliação do conteúdo com o recurso das legendas textuais, e (d) Alguns usuários surdos assistindo ao conteúdo com o avatar posicionado na tela	62
Figura 15. Nível de escolaridade dos surdos que participaram da segunda etapa do experimento	64
Figura 16. Momentos durante o teste com a solução proposta	66

Figura 17. Gráfico do percentual de compreensão do conteúdo exibido no experimento	69
Figura 18. Análise do nível de acerto entre as abordagens através do gráfico de caixa	70
Figura 19. Artigo Publicado no <i>Journal of Research and Practice in Information Technology</i> sob o tema: <i>Accessibility as a Service: Augmenting Multimedia Content with Sign Language Video</i>	81
Figura 20. Artigo publicado na II Escola Paraibana de Informática com o tema: <i>"Uma Experiência de Avaliação de Interface de Usuário em Dispositivos "Segunda Tela" no aplicativo CineLIBRAS"</i>	82
Figura 21. Prêmio de melhor artigo de iniciação científica pelo artigo intitulado como: <i>"Uma Experiência de Avaliação de Interface de Usuário em Dispositivos "Segunda Tela" no aplicativo CineLIBRAS"</i>	83

Lista de Tabelas

Tabela 1. Classificação dos tipos de perda auditiva	27
Tabela 2. Especificação das mensagens do protocolo de comunicação com o Fogo Player	49
Tabela 3. Nível de acerto das perguntas sobre o conteúdo na primeira etapa do experimento	63
Tabela 4. Classificação do nível de compreensão dos sinais realizados pelo avatar 3D	63
Tabela 5. Nível de acerto das perguntas sobre o conteúdo na segunda etapa do experimento	66
Tabela 6. Classificação do nível de compreensão dos sinais realizados pelo avatar 3D e a aderência a gramática de LIBRAS utilizando a solução proposta	67
Tabela 7. Nível de satisfação do surdo em utilizar um dispositivo móvel com a solução proposta	68

Lista de Símbolos e Abreviaturas

2D: Bidimensional

3D: Tridimensional

ABNT: Associação Brasileira de Normas Técnicas

AC: Assistente de Comunicação

ANSI: *American National Standards Institute*

ASL: *American Sign Language*

BSL: *Brazilian Sign Language*

CC: *Closed Caption*

dB: decibéis

DC: *Digital Cinema*

DCDM: *Digital Cinema Distribution Master*

DCI: *Digital Cinema Initiatives*

DGP: *Digital Cinema Package*

DCSS: *Digital Cinema System Specification*

FUNAD: Fundação Centro Integrado de Apoio à Pessoa com Deficiência

GTRI: *Georgia Tech Research Institute*

HTTP: *HyperText Transfer Protocol*

IBGE: Instituto Brasileiro de Geografia e Estatística

IrishSL: *Irish Sign Language*

ITU: *International Telecommunication Union*

JSL: *Japanese Sign Language*

LCD: *Liquid Crystal Display*

LIBRAS: Língua Brasileira de Sinais

MXF: *Material eXchange Format*

MTU: *Maximum Transmission Unit*

NBR: Norma Brasileira

OMS: Organização Mundial de Saúde

PDA: *Personal Digital Assistant*

PPM: *Prediction by Partial Matching*

PTS: *Presentation Time Stamp*

RNP: Rede Nacional de Ensino e Pesquisa

SBRC: Simpósio Brasileiro de Redes de Computadores

TIC: Tecnologia de Informação e Comunicação

TS: *Transport Stream*

TVD: Televisão Digital

UHD: *Ultra High Definition*

W3C: *World Wide Web Consortium*

WHO: *World Health Organization*

WRNP: Workshop da RNP

XML: *eXtensible Markup Language*

Sumário

1 Introdução.....	16
1.1 Motivação	18
1.2 Objetivos.....	21
1.3 Estrutura da Monografia.....	23
2 Fundamentação Teórica	25
2.1 Acessibilidade.....	25
2.2 Surdez e Deficiência Auditiva	26
2.3 Língua Brasileira de Sinais.....	28
2.4 Cinema Digital.....	31
3 Trabalhos Relacionados.....	34
3.1 Geração Automática de Conteúdos Acessíveis em Língua de Sinais	34
3.2 Conteúdos Acessíveis para Surdos em Dispositivos de Segunda Tela	36
3.2.1 Comunicação e Interação com Deficientes Auditivos em Sala de Aula	37
3.2.2 <i>Sony Glasses</i>	38
3.2.3 <i>Wireless Captioning System</i>	39
3.2.3.1 <i>New Wireless Captioning System Debuts at Dallas</i>	40
4 Solução Proposta	43
4.1 Modelo Conceitual da Infraestrutura do CineLIBRAS	43
4.2 Protocolo de Comunicação com o Fogo Player	47
4.3 Arquitetura de Componentes do CineLIBRAS	51
4.3.1 Extrator de Legendas	53
4.3.2 Gerador de LIBRAS	54
4.3.2.1 Tradutor	54
4.3.2.2 Sincronizador.....	55
4.3.3 Dicionário de LIBRAS	56

4.3.4 Distribuidor.....	57
5 Resultados	59
5.1 Teste de Inteligibilidade dos Conteúdos Gerados	59
5.1.1 Avaliação da primeira etapa do experimento	60
5.1.2 Avaliação da segunda etapa do experimento.....	64
5.1.3 Análise dos resultados obtidos nas duas etapas do experimento.....	69
6 Considerações Finais.....	72
6.1 Trabalhos Futuros	73
Referências Bibliográficas	75
ANEXO A - Questionário aplicado no experimento com os surdos.....	79
ANEXO B - Artigo Publicado no <i>Journal of Research and Practice in Information Technology</i>	81
ANEXO C - Artigo Publicado na II Escola Paraibana de Informática	82
ANEXO D - Prêmio de Melhor Artigo de Iniciação Científica na II Escola Paraibana de Informática	83

1 Introdução

O desenvolvimento inclusivo tem como objetivo promover a igualdade de oportunidades e direitos para todos, através do projeto e execução de ações políticas voltadas para o desenvolvimento socioeconômico e humano, independente de classe social, gênero, raça e das condições físicas, mentais ou sensoriais (Bieler 2005). Este, por sua vez, possui um conjunto de princípios básicos que norteiam o desenvolvimento de uma sociedade ativa e participativa, que dispõe de recursos para acessar informações, obter conhecimentos e meios para se comunicar com outras pessoas.

Para os surdos, no entanto, o atendimento a esses princípios básicos ainda não é uma realidade. Eles se deparam com diversos entraves para participar da vida em sociedade, para se comunicar com outras pessoas e para acessar informações. Sua participação em ambientes sociais como, por exemplo, cinemas, teatros, ambientes educacionais, dentre outros, é muita vezes restrita devido a ausência de soluções acessíveis que viabilizem sua participação e interação nesses ambientes.

Nos cinemas, por exemplo, os recursos de acessibilidade para os surdos, quando estão disponíveis, são limitados ao uso de legendas no formato texto. No entanto, os surdos se comunicam naturalmente através das línguas de sinais, línguas com modalidade de comunicação gestual-visual que, em virtude disso, enfrentam diversas dificuldades para ler e escrever na língua oral do seu país. Uma vez que essas línguas são baseadas em sons, muitos surdos passam vários anos na escola e não conseguem se alfabetizar na língua oral do seu país.

De acordo com o censo demográfico realizado pelo Instituto Brasileiro de Geografia e Estatística em 2000 (IBGE 2000), cerca de 97% dos surdos não conseguem concluir o ensino médio. Um outro exemplo que pode ser citado é um estudo de Wauters (2005) feito com crianças surdas holandesas de 9 a 20 anos que mostrou que apenas 25% delas possuem um nível de leitura igual ou superior ao de uma criança de 9 anos sem deficiência.

Na literatura científica, existem diversos trabalhos que visam minimizar os problemas enfrentados pelos surdos em ambientes sociais de participação coletiva (Tambassia, et al. 2012), (Kato, et al. 2010), (BUTTUSSI, et al. 2008) , (West, et al. 2009). Porém, a maioria destas propostas utilizam apenas legendas textuais, o que acaba não solucionando o problema de maneira eficaz, uma vez que as línguas orais funcionam como uma espécie de “segunda língua” para os surdos.

Este trabalho, portanto, tem como objetivo o desenvolvimento de uma solução tecnológica que viabilize a geração automática e a distribuição de vídeos em Língua Brasileira de Sinais (LIBRAS) nas salas de cinema. A proposta é que os vídeos acessíveis em LIBRAS sejam gerados automaticamente a partir das legendas do conteúdo cinematográfico e sejam transmitidos (distribuídos) e exibidos nos dispositivos móveis (*tablets, smartphones* etc.) dos usuários surdos. Assim, é possível que os surdos brasileiros possam acompanhar o conteúdo audiovisual na sua língua natural de comunicação, a LIBRAS, nesses ambientes.

Segundo os dados apresentados pela Organização Mundial de Saúde – OMS (*World Health Organization* – WHO) em 2013, pelo menos 5,3% da população mundial – correspondendo a cerca de 360 milhões de pessoas – possui algum nível de deficiência auditiva. Dentro desse mesmo universo de dados, 91% (cerca de 328 milhões) das pessoas com deficiência auditiva são adultas e 9% (cerca de 32 milhões) são crianças. E, em relação ao gênero, 56% das pessoas com deficiência auditiva são do sexo masculino, enquanto 44% são do feminino. A Figura 1 ilustra os dados divulgados pela OMS de acordo com os atributos de idade e gênero das pessoas que fizeram parte desta pesquisa.

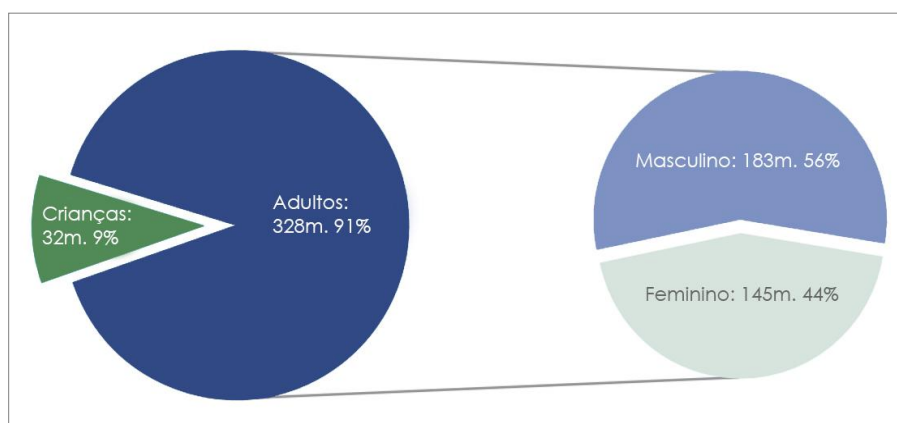


Figura 1. Deficiência auditiva: dados da população mundial (WHO 2013).

No Brasil, o número de pessoas com deficiência auditiva também é muito expressivo. De acordo com o censo demográfico feito pelo IBGE em 2010 (IBGE 2011), no Brasil existem cerca de 7,5 milhões de pessoas (aproximadamente 5,1% da população) com algum nível de deficiência auditiva. A classificação quanto ao nível de deficiência investigada foi dividida em três categorias: (1) não consegue de modo algum – 0,18%, (2) grande dificuldade – 0,94% – e (3) alguma dificuldade – 3,98%. Este resultado apontou um aumento de 1,7% em relação ao censo anterior realizado em 2000 pelo IBGE (2000).

Portanto, diante das adversidades apresentadas, que assolam boa parte da população mundial, um ponto que torna este trabalho relevante é que a partir da solução proposta, as pessoas que são portadoras de deficiências auditivas poderão participar igualmente da exibição de conteúdos cinematográficos em salas de cinema. Para isso, conforme mencionado, o propósito é dar a possibilidade ao usuário surdo de receber (através de dispositivos móveis) a tradução do filme na sua língua natural de comunicação, isto é, LIBRAS.

1.1 Motivação

Na sociedade moderna, existem diversas formas de se definir o termo acessibilidade. A mais comum está relacionada ao compromisso em melhorar a qualidade de vida das pessoas, principalmente com relação aos fatores perceptivos, cognitivos e motores (Gallahue e Ozmun 2005). Em outras palavras, acessibilidade consiste em desenvolver mecanismos, ferramentas, técnicas que possibilitem que pessoas com deficiência possam ter acesso à espaços físicos, informações, produtos e serviços, independente de deficiência.

Contudo, as pessoas com deficiência enfrentam grandes dificuldades para ter acesso a esses recursos (espaços físicos, informações, produtos e serviços). Os surdos, por exemplo, usam uma linguagem gestual-visual para se comunicar naturalmente, no entanto, isso implica em limitações de comunicação com outras pessoas, de acesso à informações, participação de ambientes educacionais, dentre outros. Além a isso, as Tecnologias da Informação e Comunicação (TIC), quando são desenvolvidas, raramente

levam em consideração os requisitos de comunicação desses usuários (Haddon & Paul 2001), o que acaba restringindo o acesso dessas pessoas no que se refere aos recursos oferecidos por essas tecnologias.

Conforme mencionado anteriormente, a maioria dos surdos possuem grandes dificuldades para participar do processo educacional e desenvolver habilidades de leitura na língua oral do seu país (Menezes e Cavalcante 2008) (IBGE 2000) (Wauters 2005). Dessa forma, soluções que utilizam o modelo da linguagem oral como meio principal de acesso à informação como, por exemplo, as legendas, tendem a restringir o acesso daqueles que possuem pelo menos o mínimo de conhecimento nesta modalidade. No entanto, tendo como base os trabalhos da literatura científica investigada, percebe-se que existem muitas propostas com a finalidade de reduzir as barreiras de comunicação e acesso à informação dos surdos (Tambassia, et al. 2012), (Kato, et al. 2010), (Ferreira, et al. 2011), (Buttussi, et al. 2008), (SONY 2012), (West, et al. 2009).

Tambassia et al. (2012) propôs uma solução baseada em reconhecimento de voz para traduzir conteúdos educacionais para LIBRAS. Essa solução reconhece a voz do educador, traduz o texto resultante para a gramática da LIBRAS e, em seguida, envia o texto e a sinalização correspondente em língua de sinais, representada usando avatar 2D para o terminal do estudante (dispositivos móveis, *desktops* ou *notebooks*). Uma das limitações dessa abordagem é que ela utiliza textos, que mesmo estando de na gramática da LIBRAS, causam certo desconforto para os surdos, por não ser a forma natural de comunicação deles. Além disso, como os sinais são naturalmente representados em 3D, um dos problemas de representá-los usando avatares 2D, é que alguns parâmetros do sinal podem ser confundidos ou identificados incorretamente. Por exemplo, duas configurações de mão podem ser confundidas, devido à oclusão nos dedos (BUTTUSSI, CHITTARO e COPPO 2007).

(Kato, et al. 2010) propuseram uma solução que realiza traduções para línguas de sinais usando intérpretes humanos localizados remotamente. A solução funciona da seguinte forma. Suponha dois ambientes, A e B. No ambiente A, onde ficam localizados os estudantes surdos que acompanharão a tradução, são instaladas duas câmeras, uma direcionada para o instrutor e outra para os participantes, e um telão onde será apresentada a tradução para LIBRAS. As imagens do ambiente A são capturadas e

enviadas para dois monitores distintos no ambiente B, onde está localizado o intérprete de língua de sinais. Em B, uma câmera captura a tradução gerada pelo intérprete de língua de sinais e a transmite para o telão LCD instalado em A. Embora utilize traduções geradas por intérpretes humanos, a primeira opção para os surdos, ela possui um alto custo operacional, que inclui equipamentos para captura e transmissão desses conteúdos, além da necessidade de intérpretes em tempo integral.

No contexto da TV Digital (TVD), Ferreira et al. (2011) propuseram uma solução que utiliza os recursos da TVD para fornecer suporte à geração automática de janelas em Língua Brasileira de Sinais (LIBRAS) a partir da tradução automática do *Closed Caption* (CC) transmitido pela estação de TV. A proposta tem como objetivo prover suporte para línguas de sinais em *middlewares* de TVD compatíveis com a especificação *International Telecommunication Union – ITU J.202*, utilizando apenas as *Application Programming Interface* (API) e os componentes definidos nessa especificação.

Buttussi et al. (2008) propuseram uma solução que tinha como objetivo facilitar a comunicação entre médicos e pacientes surdos, especialmente em soluções que envolvem emergência médica, onde a comunicação médico-paciente é, muitas vezes, essencial para o diagnóstico e tratamento do problema. Esta solução utiliza dispositivos móveis, como, por exemplo, *tablets*, *smartphones* e PDAs para lidar com esses problemas. Sua estratégia consiste em apresentar sentenças simplificadas sobre o estado do paciente. As sentenças são organizadas de forma hierárquica, separadas por grupos (questões primordiais, questões sobre sintomas, questões sobre problemas cardiovasculares etc.), formuladas de forma que a resposta do paciente se limite a sim/não, ou um número correspondente. Dessa forma, a medida em que o médico seleciona as sentenças (perguntas), elas são apresentadas no *display* do dispositivo em língua de sinais, facilitando a compreensão dos pacientes surdos. Na avaliação subjetiva da solução, 80% dos médicos afirmaram que o sistema ajudou de alguma forma, 10% afirmaram que a solução ajudou muito no diagnóstico e tratamento, enquanto que 10% deles responderam que ela foi indiferente.

Pesquisadores da *Georgia Tech Research Institute* (GTRI) desenvolveram um sistema de legenda pessoal para diferentes ambientes. O objetivo é disponibilizar

legendas para as pessoas nos cinemas, arenas de esporte, escolas, teatros, reuniões governamentais, museus, igrejas, dentre outros. West et al. (2009) ressalta que o sistema utiliza o protocolo *wireless* 802.11b e que consiste basicamente de três componentes: (1) transmissor: responsável pela distribuição das legendas (pré-gravada, digitada em tempo real ou por reconhecimento de voz); (2) receptor: dispositivo móvel (PDA, *smartphones* e laptops) equipado com interface de rede para receber as legendas; (3) *display*: tela onde as legendas são exibidas.

De acordo com a especificação do GTRI, a apresentação das legendas pode ser na própria tela do dispositivo móvel ou em *microdisplays* conectados ao dispositivo. Na versão do *microdisplay*, são utilizados óculos (monocular ou binocular) que criam a ilusão de que a legenda está sobreposta no campo de visão do usuário (West, et al. 2009). Entretanto, conforme mencionado anteriormente, essa é uma abordagem que atribui certa limitação para a comunidade surda de modo geral, pois como sua língua natural de comunicação é baseada em gestos e expressões faciais, torna-se imprescindível o domínio da língua oral para compreender as legendas no formato texto.

Contudo, baseado na literatura apresentada, nenhuma das soluções investigadas apresentam meios para viabilizar a geração automática e distribuição de janelas em Línguas de Sinais (LS) através de um agente virtual 3D no Cinema Digital. Portanto, esse é o embasamento principal que motiva a pesquisa e o desenvolvimento de uma solução nesse contexto. Mais especificamente, a motivação dessa proposta é reduzir as barreiras de acesso à informação dos surdos nas salas de cinema, desenvolvendo uma solução tecnológica que permita o acompanhamento da tradução do conteúdo cinematográfico na sua língua natural de comunicação, a LIBRAS.

1.2 Objetivos

O objetivo geral deste trabalho é desenvolver uma solução computacional, denominada CineLIBRAS, que permita a geração automática e distribuição de vídeos em LIBRAS, para usuários com baixa audição, nas salas de cinema. Conforme mencionando anteriormente, a proposta é que vídeos de LIBRAS sejam gerados automaticamente a

partir da tradução da legenda do conteúdo cinematográfico e que sejam distribuídos para dispositivos móveis dos usuários surdos. Com isso, é possível que eles acompanhem a tradução dos conteúdos cinematográficos para LIBRAS, sua língua natural de comunicação, em um dispositivo de segunda tela, tornando a sessão do cinema inclusiva.

O projeto arquitetural da proposta está fundamentado em três vertentes: (1) adaptabilidade, a solução proposta deve adaptar-se facilmente ao cenário do Cinema Digital; (2) flexibilidade, o solução deve permitir que os usuários possam se conectar após o início da exibição do conteúdo, mantendo-os sempre sincronizados, de acordo com a taxa de atraso permitida; (3) interoperabilidade, a solução deve atender as demandas independente da plataforma em que os clientes (dispositivos móveis) estejam executando.

Para alcançar o objetivo geral desta proposta, foram definidos os seguintes objetivos específicos:

- Objetivo 1.* Elaboração de um instrumento de avaliação que permita identificar as principais necessidades relacionadas a acessibilidade em ambientes cinematográficos;
- Objetivo 2.* Definição de um protocolo de comunicação entre o *player* de Cinema Digital e o sistema proposto para o recebimento dos arquivos de legenda;
- Objetivo 3.* Definição de uma arquitetura geral do sistema que inclua a comunicação com o *player* de Cinema Digital, os componentes de extração, tradução e disponibilização dos conteúdos acessíveis até o recebimento do conteúdo pelos dispositivos móveis;
- Objetivo 4.* Investigação de uma estratégia de apresentação do vídeo acessível em língua de sinais nos dispositivos móveis de forma sincronizada com o conteúdo principal;
- Objetivo 5.* Desenvolvimento de um protótipo do sistema e sua aplicação em cenários de testes com usuários surdos.

1.3 Estrutura da Monografia

A estrutura deste trabalho está organizada da seguinte forma: no primeiro capítulo foi apresentada uma contextualização do problema de pesquisa, ressaltando a dificuldade de comunicação e acesso à informação dos surdos, especialmente em ambientes públicos, educacionais e nas TICs. O segundo capítulo apresenta uma fundamentação teórica dos assuntos inerentes a linha de pesquisa desse trabalho. Neste, alguns conceitos-chave como acessibilidade, deficiência auditiva, Língua Brasileira de Sinais (LIBRAS) e Cinemas Digitais serão apresentados e discutidos com mais detalhes.

O terceiro capítulo apresenta alguns trabalhos que foram selecionados de acordo com o grau de relevância com a solução proposta. Esse capítulo tem o objetivo de ressaltar as principais contribuições da literatura científica no que se refere a redução da problemática enfrentada pelos deficientes auditivos em alguns cenários como, por exemplo, em ambientes sociais de participação coletiva (cinemas, teatros, estádios de futebol) e no âmbito educacional.

O quarto capítulo apresenta as características inerentes ao desenvolvimento da solução proposta neste trabalho. Nesse capítulo, é exposto o modelo conceitual da infraestrutura da solução, assim como a especificação do protocolo de comunicação com o sistema de distribuição de conteúdos cinematográficos e a arquitetura dos componentes de software desenvolvidos para essa solução.

O quinto capítulo mostra alguns resultados obtidos após um experimento realizado com um protótipo da solução proposta. O objetivo desse capítulo é apresentar o nível de compreensão dos conteúdos gerados pela solução. Além disso, esse capítulo apresenta um comparativo entre os resultados dessa solução e os de outras abordagens investigadas.

Por fim, o sexto capítulo expõe as considerações finais. Neste, são discutidos os aspectos gerais da proposta com relação aos objetivos definidos. Esses aspectos ressaltam as limitações encontradas durante o desenvolvimento deste trabalho e, com o

objetivo de sumarizar os principais pontos sobre o assunto, apresenta algumas sugestões de pesquisas para a elaboração de trabalhos futuros.

2 Fundamentação Teórica

Neste trabalho, alguns conceitos-chave como acessibilidade, deficiência auditiva, língua brasileira de sinais (LIBRAS) e cinemas digitais são recorrentemente utilizados e fundamentam o trabalho. Em consequência disso, neste capítulo, serão apresentados e descritos os principais conceitos, propriedades e características relacionados a esses temas acima mencionados.

2.1 Acessibilidade

Segundo a Associação Brasileira de Normas Técnicas (ABNT NBR 9050), o termo acessibilidade é definido como a possibilidade e condição de alcance, percepção e entendimento para a utilização com segurança e autonomia de edificações, espaço, mobiliário, equipamento urbano e elementos.

O vocábulo acessibilidade significa incluir pessoas com deficiência na participação de atividades como o uso de produtos, serviços e informações (BRASIL, Acessibilidade 2011). Para (Amaral e Souza s.d.), o acesso à informação é algo complexo, pois implicam diretamente na constituição dos direitos humanos. Segundo a Lei Nº 10.098, de 19 de Dezembro de 2000 (BRASIL, Lei n. 10.098/2000 2000) acessibilidade é:

"possibilidade e condição de alcance para utilização, com segurança e autonomia, dos espaços, mobiliários e equipamentos urbanos, das edificações, dos transportes e dos sistemas e meios de comunicação, por pessoa portadora de deficiência ou com mobilidade reduzida."

Os direitos conferidos as pessoas com necessidades especiais não devem ser desprezados nem suprimidos, pois, os benefícios ostentados pelas inovações tecnológicas (comunicação, informação, entretenimento etc.) quando combinados com as habilidades desses indivíduos, podem agregar grandes valores para a sociedade de modo geral.

Segundo (Pupo, Melo e Ferrés 2006), no entanto, atualmente, é possível perceber uma considerável discrepância entre o grande avanço tecnológico e os tímidos avanços sociais. Essa é uma lacuna que necessita ser preenchida por pessoas que acreditam na inclusão como ruptura dos paradigmas existentes, ressalta (Pupo, Melo e Ferrés 2006) para não deixar ninguém de fora na construção de ambientes acessíveis.

2.2 Surdez e Deficiência Auditiva

Deficiência auditiva, segundo (FEBRABAN 2010) é aquela deficiência que, por motivo de perda ou anomalia congênita ou adquirida, parcial ou total, de estrutura ou função da audição, pode ocasionar restrições da capacidade de comunicação, de interpretação sobre as condições de segurança e de orientação.

De acordo com o Decreto de lei nº 5.296 (2004) da constituição brasileira, deficiência auditiva consiste na perda bilateral, parcial ou total, de quarenta e um decibéis (dB) ou mais, aferida por audiograma nas frequências de 500Hz, 1.000Hz, 2.000Hz e 3.000Hz. A insuficiência do canal auditivo pode ser compreendida de forma genérica como o contraste existente entre a performance do indivíduo e a habilidade normal de percepção sonora de acordo com os padrões estabelecidos pela *American National Standards Institute* (ANSI 1989).

A natureza da surdez, muitas vezes confundida ou mal interpretada, não implica na perda total da capacidade de percepção dos sons. Embora existam indivíduos que são completamente surdos, há também casos em que a lesão do aparelho auditivo não compromete a audição literalmente. Nesse caso, é bastante comum afirmar que o indivíduo possui perda auditiva funcional. De acordo com (Brasilmedia 2012), a perda auditiva funcional pode ser classificada nas seguintes categorias:

- **Leve:** o indivíduo é incapaz de perceber a emissão de sons reproduzidos com frequência abaixo de 30 dB. Nessas condições, participar de discursos, por exemplo, pode resultar em situações frustrantes, principalmente se o ambiente dispor de muitos ruídos de fundo;

- **Moderada:** nesta categoria, frequências sonoras emitidas abaixo de 50 dB tendem a não serem perceptíveis. Dependendo da ocasião, para que não ocorram perdas de informações e possíveis circunstâncias de desconforto, aparelho ou prótese auditiva pode ser necessário;
- **Severa:** o indivíduo não está hábil para perceber ruídos sonoros abaixo de 80 dB. Em algumas situações, os aparelhos e próteses auditivas resolvem o problema, mas nem sempre são eficazes. Alguns indivíduos com perda auditiva severa se comunicam através de linguagem gestual, outros contam com técnicas de leitura labial;
- **Profunda:** nenhuma sensação auditiva pode ser captada espontaneamente quando estiver com frequência abaixo de 95 dB. Os indivíduos que se encontram nesta categoria devem recorrer à linguagem gestual para se comunicarem e obterem informações ou, desenvolverem habilidades de leitura labial.

Para (CFRF 2009), a classificação do tipo de perda auditiva leva em consideração a comparação dos limiares entre a via aérea e a via óssea de cada orelha e, portanto, é imprescindível realizar a pesquisa dos limiares tonais por via aérea e óssea. A Tabela 1 apresenta de forma sucinta a relação entre as características inerentes a cada tipo de perda auditiva.

Tabela 1. Classificação dos tipos de perda auditiva.

Tipo	Características
Perda auditiva condutiva	Avaria localizada na região externa ou intermediária do composto auditivo. Quando ocorrem lesões nessas regiões, os ossículos (martelo, bigorna e estribo) podem não vibrar adequadamente, impedindo a detecção das informações sonoras.
Perda auditiva neurossensorial (ou sensorio-neural)	Problemas na região interna da estrutura auditiva ou, no sistema auditivo periférico, cientificamente denominado patologia coclear e retrococlear. Nesse caso, os ossos da parte interna podem vibrar perfeitamente, mas os

	nervos auditivos responsáveis são incapazes de transmitir informações sonoras para o cérebro.
Perda auditiva mista	A perda auditiva mista abrange tanto a perda condutiva quanto a neurossensorial. Esta é comumente diagnosticada perda auditiva por apresentar problemas nas regiões externas, intermediária e interna do composto auditivo. O tratamento para esses casos pode ser feito através de procedimentos cirúrgicos ou, com o auxílio de aparelhos ou próteses auditivas.

2.3 Língua Brasileira de Sinais

A comunicação faz parte do cotidiano de todo ser humano. Ela está presente em grande parte das atividades que realizamos, seja conversando com os amigos, assistindo algum programa de entretenimento, em gestos de desaprovação no trânsito e muitos outros. Dessa forma, podemos concluir que a comunicação é um fator extremamente essencial para a vida das pessoas. Ou seja, é praticamente impossível desassociar nossa vida das nossas necessidades de comunicação.

Para uma comunicação efetiva, é necessário que a mesma seja clara e livre de obstruções, pois havendo ruídos no canal de transmissão, a informação transmitida pode ter sua integridade comprometida, perdida ou até mesmo impossibilitada. Ao longo da vida, segundo (UNESCO 2010), a educação baseia-se em quatro pilares: aprender a conhecer, aprender a fazer, aprender a conviver e aprender a ser. Para o Instituto Passadori (2012), a melhoria do processo educacional será possível apenas se a capacidade de comunicação dos elementos que compõem a rede da Educação (família, escola, mídias e sociedade), estiverem desenvolvidas e preparadas para cumprir seu papel.

Da mesma forma como cada país utiliza um determinado idioma para que as pessoas possam se comunicar, por exemplo, português no Brasil, inglês nos Estados Unidos, árabe no Oriente Médio etc., pessoas com deficiência auditiva, no mundo

inteiro, utilizam uma língua específica para se comunicar, as línguas de sinais. Essas línguas diferem claramente das línguas orais citadas anteriormente. A combinação de movimentos utilizando principalmente os membros superiores do corpo (mãos, braços, cabeça, tronco) permite que o emissor expresse diversos conceitos como, por exemplo, descritivo, emotivo, literal, metafórico, concreto e abstrato, e que o receptor perceba e compreenda esses conceitos utilizando o sistema visual (D. A. Silva 2012).

Segundo a Lei Brasileira No 10.436, de 24 de Abril de 2002, a língua de sinais oficial do Brasil, e utilizada pela maioria dos surdos brasileiros, é a Língua Brasileira de Sinais - LIBRAS. A LIBRAS, assim como as outras línguas de sinais, possuem uma estrutura gramatical própria, composta pelos diversos níveis linguísticos como morfologia, sintaxe e semântica. De forma similar às línguas orais, elas também possuem itens léxicos que são denominados sinais.

Os sinais são compostos por fonemas que são as unidades básicas da comunicação baseada em sinais. Segundo Buttussi et al. (2007), um sinal consiste e é unicamente identificado por cinco fonemas:

1. **Configuração da mão:** representa a posição dos dedos e seus movimentos. Um sinal pode ser produzido pela mão predominante (mão direita para os destros) ou pelas duas mãos. Uma configuração de mão pode se diferenciar das demais pela extensão (lugar e número de dedos estendidos), pela contração (mão aberta ou fechada), e pelo contato ou divergência dos dedos. Segundo Felipe (2007), em LIBRAS existem atualmente 60 possíveis configurações de mão, dentre as quais 26 são usadas para representar as letras do alfabeto (ver Figura 2).
2. **Ponto de Articulação:** representa a parte do corpo do emissor onde os sinais são realizados (ou começam a ser realizados). Em LIBRAS, esse ponto de articulação pode estar localizado em alguma parte do corpo ou estar localizado em um espaço neutro vertical (do meio do corpo até a cabeça) e horizontal (à frente do emissor). Por exemplo, os sinais TRABALHAR, BRINCAR, CONSERTAR são executados no espaço neutro, enquanto que os sinais ESQUECER, APRENDER e PENSAR são feitos na testa.

3. **Movimento:** um sinal pode ter movimentos ou não. O movimento representa o deslocamento da mão no espaço durante a execução do sinal. Caracterizar os movimentos não é uma tarefa trivial, uma vez que a mão é um objeto assimétrico e que os eixos podem se deslocar simultaneamente. Segundo Fusco (2004), no entanto, a maioria dos sinais pode ser dividida em pequenos segmentos de movimentos e cada um dos quais relacionados a um eixo. Em LIBRAS, por exemplo, os sinais AJOELHAR e EM-PÉ não tem movimento.
4. **Orientação:** representa a direção ou orientação do movimento. Os sinais que possuem algum tipo de movimento são executados em uma determinada direção. Em LIBRAS, a inversão da direção pode representar a ideia de oposto, contrário ou modificar a concordância número-pessoal do sinal.
5. **Expressões não manuais:** muitos sinais podem requerer características adicionais para expressar sentimentos de alegria, de tristeza, uma pergunta ou uma exclamação e necessitam de um traço diferenciador, como uma expressão facial e/ou corporal, para poder expressar essas características. Dessa forma, as expressões não manuais que podem ser expressões faciais e/ou corporais podem assumir tanto uma função léxica (expressar alegria e tristeza, por exemplo) quanto uma função sintática (expressar uma pergunta ou uma exclamação, por exemplo) na estrutura dos sinais. A inclusão de uma expressão facial e/ou corporal no discurso pode alterar o significado de um sinal e sua ausência pode produzir um sinal sem sentido. Alguns sinais em LIBRAS, inclusive, são feitos apenas com a bochecha, como LADRÃO e ATO-SEXUAL.

Conforme Buttussi et al. (2007), a partir da combinação desses cinco fonemas formam-se os sinais, que permitem comunicar diferentes ideias ou aspectos de uma sentença. Expressar-se em LIBRAS ou em qualquer outra língua de sinais, portanto, é combinar esses elementos para formar os sinais e combinar os sinais para formar sentenças em um determinado contexto.

Assim como nas línguas de modalidade oral, as línguas de sinais possuem um conjunto de sinais para representar o alfabeto. Em LIBRAS, por exemplo, para cada sinal do alfabeto existe um gesto com as mãos. O alfabeto manual das línguas de sinais

teve origem pela necessidade de representar as letras de forma visual e era usado principalmente para ensinar pessoas surdas a ler e escrever (Silva, et al. 2007). A Figura 2 ilustra as configurações de cada sinal que constitui o alfabeto da Língua Brasileira de Sinais (LIBRAS).



Figura 2. Alfabeto da Língua Brasileira de Sinais.

2.4 Cinema Digital

O Cinema é um ambiente capaz de ampliar a capacidade perceptiva do pensamento humano, mobilizando não apenas o olhar do espectador, mas o corpo em sua totalidade (Gonçalves 2008). Nos primórdios da produção cinematográfica, sua realização não passava da predominância de um conjunto de imagens, na qual engloba não apenas a história das práticas de projeção de imagens, mas também a dos divertimentos populares, dos instrumentos óticos e das pesquisas com fotografias (Mascarello 2006).

Contudo, em virtude dos avanços tecnológicos, a indústria cinematográfica vêm investindo cada vez mais no desenvolvimento de tecnologias que possibilitam a produção, distribuição e reprodução de filmes digitais em um nível de qualidade equivalente aos filmes analógicos produzidos com as películas de 35mm (DCI 2012).

Segundo a *Digital Cinema Initiatives* (DCI 2012), essas tecnologias permitem a digitalização de filmes em alta resolução, compressão de imagem digital, redes de computadores com alta capacidade para armazenamento de dados, a projeção digital avançada, dentre outros. A combinação desses recursos deram origem a um novo conceito no cenário do cinema e da tecnologia, o Cinema Digital.

Para Ferreira et al. (2002), o Cinema Digital proporcionou mudanças significativas no processo de produção dos conteúdos audiovisuais. As vantagens advindas do Cinema Digital em relação ao modelo analógico (isto é, baseado em películas) trouxeram benefícios concebíveis às indústrias cinematográficas, com maior destaque nos seguintes aspectos:

- **Custo:** as películas utilizadas no cinema analógico possuem um custo superior em relação aos recursos digitais;
- **Flexibilidade:** a manipulação e edição dos conteúdos digitais tornou-se relativamente mais simples para o produtor cinematográfico;
- **Distribuição:** a redução nos custos de distribuição são extremamente significativos, sendo quase que indiferente monetariamente exibir um filme em 10 cinemas ou em 1000, pois, em vez de ter de copiar o conteúdo em várias películas, o conteúdo poderá ser gravado em suportes físicos flexíveis, como o DVD-ROM, difundidos por cabo ou até mesmo por satélite.

A concepção dos padrões técnicos de produção dos conteúdos digitais no cinema foram elaborados pela *Digital Cinema Initiatives* (DCI 2012). A DCI é uma entidade que foi criada por um conjunto formado por sete empresas do ramo da produção cinematográfica, são elas: *Disney*, *Fox*, *Metro-Goldwyn-Mayer*, *Paramount Pictures*, *Sony Pictures Entertainment*, *Universal Studios*, e *Warner Bros. Studios*. O envolvimento dessas empresas foi importante para garantir o máximo de interoperabilidade e compatibilidade com os produtos e serviços produzidos pelas demais indústrias participantes desse mercado.

Segundo a DCI (2012), as diretrizes que constituem as especificações e requerimentos necessários para a masterização, distribuição e reprodução dos conteúdos digitais no cinema está dividido da seguinte forma:

- **Digital Cinema Distribution Master (DCDM):** fornece as especificações para a manipulação de imagens, áudio e legendas (texto com marcações de tempo e imagens como legenda);
- **Compression (Image):** especifica a compatibilidade do DCI com o *codestream* JPEG 2000 e o decodificador JPEG 2000;
- **Packaging:** define os requerimentos para o empacotamento dos arquivos do DCDM nos formatos *Material eXchange Format* (MXF) e *eXtensible Markup Language* (XML). A saída desse processo é o *Digital Cinema Package* (DCP). Esta seção também define os requerimentos para a criptografia de imagens, sons e legendas do DCP;
- **Transport:** define a forma de distribuição dos conteúdos para os ambientes de reprodução usando mídias de natureza física, redes virtuais privadas ou comunicações via satélite;
- **Theater Systems:** fornece todos os requerimentos necessários para a reprodução dos conteúdos nas salas de cinema. Isso engloba projetores digitais, blocos de mídia, sistemas de armazenamento, sistemas de som, os arquivos DCP, automação do ambiente, sistema de gestão da tela e sistemas de gestão teatro (cinema).
- **Projection:** define o padrão de projeção, além do ambiente de controle, juntamente com as modificações permitidas nos parâmetros de imagem críticas para dominar e aplicações Exposição Geral.
- **Security:** fornece os requerimentos e especificações fundamentais para a persistência dos conteúdos de projeção e acesso controlado em uma arquitetura de segurança aberta.

3 Trabalhos Relacionados

Este capítulo apresenta alguns trabalhos relacionados com a solução proposta. Esses trabalhos foram selecionados de maneira independente, ou seja, não houve uma definição de critérios para a escolha dos trabalhos, sendo feita apenas análises subjetivas das contribuições. As características tomadas como base para o processo de investigação dos trabalhos foram divididos em duas vertentes: a primeira aborda assuntos relacionados a geração automática de conteúdos acessíveis em Língua de Sinais, isto é, tradução automática da língua oral para a língua de sinais (Gallo, et al. 2009) (Morrissey 2008) (Othman e Jemni 2011) (San-segundo, et al. 2011) (Veale, Collins e Conway 1998) (Zhao, et al. 2000) (Halawani 2008).

Na segunda, são apresentados alguns trabalhos relacionados com uma perspectiva voltada para a utilização de dispositivos móveis como principal meio de acesso à informação. Neste, o propósito é mostrar algumas contribuições relevantes que abordam o conceito de geração automática de conteúdos acessíveis, mas, com foco em acessibilidade para deficientes auditivos a partir da utilização de dispositivos de segunda tela (Tambassia, et al. 2012) (SONY 2012) (West, et al. 2009).

3.1 Geração Automática de Conteúdos Acessíveis em Língua de Sinais

Existe uma gama de trabalhos na literatura científica que tem como objetivo a diminuição dos problemas de comunicação enfrentados pelos deficientes auditivos (Gallo, et al. 2009) (Morrissey 2008) (Othman e Jemni 2011) (San-segundo, et al. 2011) (Veale, Collins e Conway 1998) (Zhao, et al. 2000) (Halawani 2008). Uma das estratégias investigadas nesses trabalhos consiste na tradução automática de conteúdos baseados na língua oral para a língua de sinais.

Assim como as línguas orais definem estruturas gramaticais próprias, com regras específicas em seus níveis linguísticos, morfológico e sintático, as línguas de sinais possuem um conjunto de regras gramaticais completamente diferente. Dessa forma, a tradução de conteúdos entre essas duas linguagens (por exemplo, da língua portuguesa

para LIBRAS) sem a utilização de um pré-processamento que seja capaz de analisar o contexto, ou até mesmo interpretar a ideia da língua de origem, é uma atividade propensa ao improvável (Araújo 2012).

Veale et al. (1998), por exemplo, propôs um sistema de tradução automática multilíngue para traduzir textos em inglês para língua americana de sinais (*American Sign Language* - ASL), língua irlandesa de sinais (*Irish Sign Language* - IrishSL) e língua japonesa de sinais (*Japanese Sign Language* - JSL). Esse sistema é baseado numa arquitetura de quadro negro (*blackboard control architecture*) (Othman e Jemni 2011) e possui um conjunto de agentes que cooperam para gerar os conteúdos traduzidos. Esse trabalho explora e estende alguns conceitos de Inteligência Artificial (IA) para línguas de sinais como, por exemplo, representação do conhecimento, raciocínio metafórico, arquiteturas baseadas em quadro negro (Morrissey 2008), mas nenhum teste ou experimento foi realizado para avaliar a viabilidade e qualidade solução.

Zhao et al. (2000) propuseram uma abordagem baseada em interlíngua para tradução de textos em inglês para ASL. Nessa solução, os dados de entrada são analisados e uma representação intermediária (*Intermediate Representation* - IR) é gerada a partir da análise desses dados. Um sintetizador então utiliza essa representação intermediária para gerar os sinais. No entanto, da mesma forma, que no trabalho proposto por Veale et al. (1998), nenhum teste ou experimento foi realizado para avaliar a solução. Othman & Jemni (2011) propuseram uma estratégia para alinhamento de palavras e incluíram-na em tradutor estatístico de inglês para ASL. No entanto, apenas a estratégia de alinhamento de palavras foi avaliada, não sendo realizado nenhum teste para avaliar a qualidade e velocidade da tradução, por exemplo.

Gallo et al. (2009) e San-segundo et al. (2011) propuseram uma arquitetura para traduzir voz em língua espanhola de sinais (LSE) com enfoque em ajudar pessoas surdas no atendimento em serviços públicos como, por exemplo, na renovação da carteira nacional de habilitação. A ideia do sistema é facilitar o diálogo entre surdos e prestadores de serviços públicos nesse tipo de serviço. No entanto, essa solução também é restrita a um domínio de aplicação específico (atendimento em serviços públicos). Além disso, o tempo médio reportado para traduzir cada sentença de voz para LSE foi

de cerca de 8 segundos, o que torna a solução inviável em cenários que exigem tradução em tempo real, como, por exemplo, na TV.

3.2 Conteúdos Acessíveis para Surdos em Dispositivos de Segunda Tela

Estudos empíricos na área de acessibilidade têm mostrado que as pessoas com deficiência auditiva, seja ela parcial ou total, encontram muitos obstáculos para ter acesso à ambientes sociais como, por exemplo, nos cinemas, teatros, ambientes educacionais, dentre outros (Tambassia, et al. 2012), (Kato, et al. 2010), (Buttussi, et al. 2008), (SONY 2012), (West, et al. 2009). Esses problemas acontecem principalmente pelo déficit de soluções tecnológicas que atendam suas necessidades de comunicação.

Atualmente, existem três tipos principais de estratégias para solucionar esse problema. A primeira consiste na transcrição textual do conteúdo do áudio, denominada legenda (Tambassia, et al. 2012), (West, et al. 2009), (SONY 2012). Todavia, de acordo com (Lima 2006), os surdos tem dificuldade para ler e escrever na língua oral do seu país. A segunda utiliza intérpretes de línguas de sinais para traduzir os conteúdos transmitidos na língua oral. No entanto, embora a segunda estratégia seja a solução preferida dos surdos, em geral, essa abordagem possui um alto custo operacional envolvido, pois necessita de intérpretes disponíveis em tempo integral. Além do mais, em alguns cenários essa estratégia é totalmente inviável como, por exemplo, na Web, dada a natureza dinâmica dos seus conteúdos (Filho e Araújo 2011).

A terceira estratégia envolve o uso de soluções baseadas na tradução automática de fluxos de áudio ou texto para vídeos em línguas de sinais (Tambassia, et al. 2012), (Ferreira, et al. 2011), (Filho e Araújo 2011). Nesse tipo de solução, os vídeos em línguas de sinais são, em geral, representados por avatares 3D (agentes virtuais), que exercem o papel do intérprete de língua de sinais no processo de comunicação. Esse tipo de solução não é a preferida pelos surdos, uma vez que a sinalização virtual ainda não é tão natural quanto a sinalização humana, mas é uma solução prática e viável, especialmente quando intérpretes de línguas de sinais não estão disponíveis ou não são viáveis. Conforme será apresentado no Capítulo 4, a solução proposta neste trabalho baseia-se neste tipo de abordagem.

As próximas subseções apresentam alguns trabalhos presentes na literatura científica que foram selecionados de maneira independente (isto é, utilizando apenas a análise subjetiva como critério de seleção), mas que estão diretamente correlacionados com os objetivos deste trabalho.

3.2.1 Comunicação e Interação com Deficientes Auditivos em Sala de Aula

Para viabilizar a inclusão dos surdos em escolas comuns, é necessário adequar a metodologia de ensino e os materiais didáticos para permitir o aprendizado e interação dessas pessoas (Tambassia, et al. 2012).

Com base nessa perspectiva, Tambassia et al. (2012) propôs uma solução que utiliza mecanismos de reconhecimento de voz e tradução automática para traduzir conteúdos educacionais para línguas de sinais. A ideia é equipar o professor (instrutor) com um microfone, responsável por capturar as informações auditivas produzidas pelo professor durante a aula. Esse áudio passa então por um processo de reconhecimento de voz, sendo convertido em formato texto, e, em seguida, o texto é adaptado para a gramática da LIBRAS, através da aplicação de um conjunto de regras de tradução.

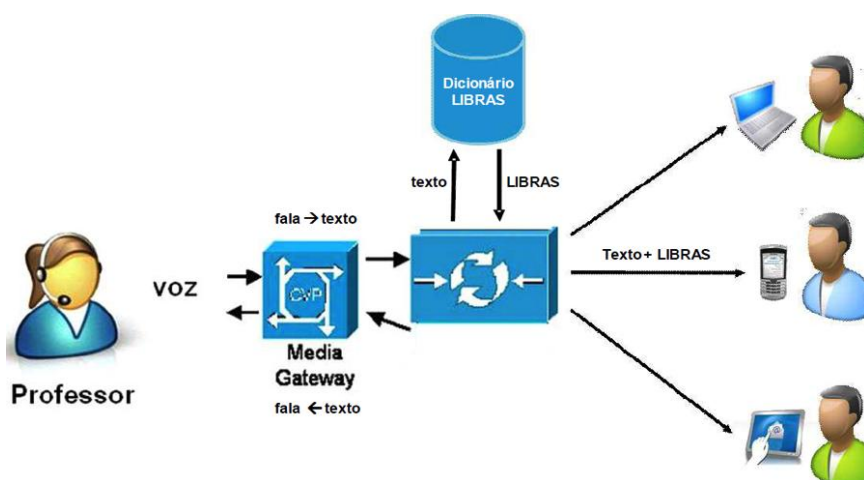


Figura 3. Arquitetura proposta por Tambassia et al. (2012).

A solução também prevê a possibilidade de apresentar os conteúdos em LIBRAS. Para isso, foi produzido um dicionário em LIBRAS, em que os sinais são representados por um avatar 2D, de acordo com o contexto de cada disciplina. O conteúdo gerado em LIBRAS, a partir desse dicionário, é então transmitido em uma rede local, juntamente com o texto, para os dispositivos móveis dos estudantes. A Figura 3 apresenta a arquitetura proposta por (Tambassia, et al. 2012).

Conforme mencionado no Capítulo 1, uma das limitações desse trabalho é que os sinais são representados usando avatares 2D. Como os sinais em LIBRAS são representados naturalmente em 3D, um dos problemas de representá-los usando avatares 2D é que alguns parâmetros do sinal podem ser confundidos ou identificados incorretamente. Por exemplo, duas configurações de mão podem ser confundidas, devido à oclusão nos dedos (BUTTUSSI, CHITTARO e COPPO 2007).

3.2.2 Sony Glasses

O cinema tem passado por grandes mudanças nos últimos anos. Essas mudanças envolvem a melhoria na qualidade dos conteúdos audiovisuais produzidos, a inclusão de tecnologias para exibição de conteúdos em 3D e em *Ultra High Definition* (UHD), dentre outros.

No entanto, essa evolução tecnológica nos Cinemas não vem acompanhada da melhoria nas técnicas de acessibilidade para deficientes cognitivos. Com o intuito de reverter essa situação, (SONY 2012) propôs uma solução que envolve o uso de óculos especiais para a exibição de legendas e de conteúdos audiodescritos.

A proposta é utilizar óculos capazes de exibir as legendas do conteúdo cinematográfico em suas lentes (*microdisplays*). Eles também estão equipados com equipamento com recursos de áudio (STW-C140GI) e, como um transmissor de dados (STWA-C101). Os recursos de áudio são importantes porque podem permitir que conteúdos audiodescritos possam ser transmitidos para cegos, facilitando também a sua inclusão nesses ambientes.

Segundo os autores, com suas tecnologias exclusivas de holografia, a projeção das legendas parecem flutuar no espaço, ou seja, o usuário tem a sensação de que o texto está sendo projetado em terceira dimensão, próximo ao *display* de exibição do conteúdo. A Figura 4 mostra a visão esquemática da sua solução.

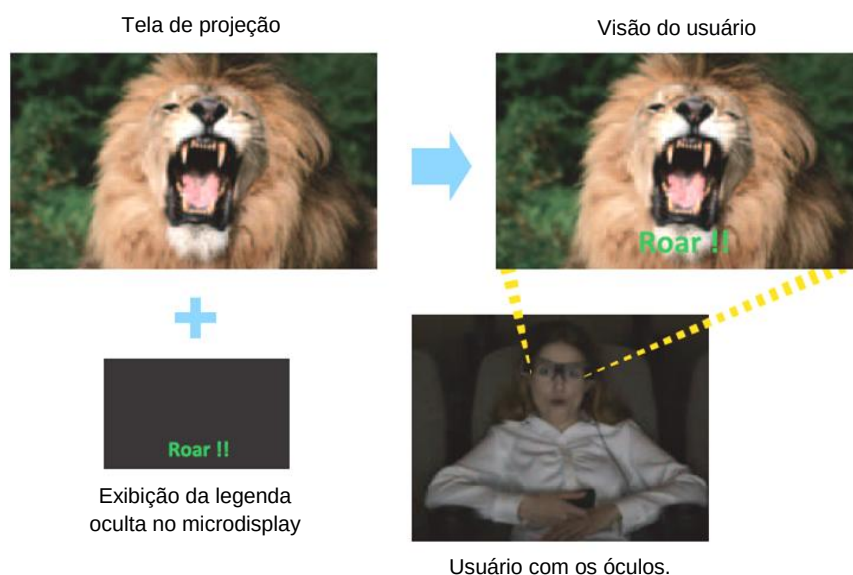


Figura 4. Visão esquemática da solução da SONY (2012).

3.2.3 Wireless Captioning System

Pesquisadores da *Georgia Tech Research Institute* (GTRI) desenvolveram um sistema de legenda pessoal com o objetivo de fornecer informações e, consequentemente, promover a acessibilidade em ambientes públicos. A proposta consiste na distribuição em larga escala (isto é, para uma grande quantidade de pessoas) do conteúdo presente nas legendas. A ideia é que esses conteúdos estejam disponíveis em ambientes sociais de participação coletiva como, por exemplo, cinemas, arenas de esporte, instituições de ensino, teatros, reuniões governamentais, museus, igrejas, dentre outros.

O propósito de West et al. (2009) é desenvolver um Assistente de Comunicação (AC) utilizando o protocolo *wireless* 802.11b e tecnologias com *microdisplay* para a projeção dos conteúdos acessíveis (isto é, legenda no formato texto). Para isso, são

disponibilizados dispositivos óticos (isto é, óculos) que podem ter uma função binocular ou monocular, onde binocular utiliza as duas lentes (ou *microdisplay*) do óculos, enquanto a monocular usa apenas uma.

West et al. (2009) ressalta que, embora posicionado próximo ao olho, o *microdisplay* utiliza lentes que fazem com que o conteúdo pareça estar flutuado a vários metros de distância com relação ao usuário, assim, é possível proporcionar uma visualização suave de ambas as imagens (isto é, tela de reprodução do filme original e o conteúdo da legenda), assim como a capacidade de percepção do mundo ao redor.

Segundo West et al. (2009), a arquitetura desta solução é composta basicamente por três componentes, são eles:

1. *Transmissor*: componente responsável pelo envio (distribuição) das legendas, que podem ser pré-gravadas, digitadas em tempo real ou geradas a partir de reconhecimento de voz;
2. *Receptor*: dispositivos móveis (PDA, *smartphones* e laptops) equipados com interface de rede com suporte ao protocolo 802.11b, para receber o conteúdo das legendas;
3. *Display*: interface de exibição das legendas, que de acordo com a especificação do GTRI, pode ser a própria tela do dispositivo móvel ou um *microdisplay* (por exemplo, óculos) conectado ao dispositivo.

3.2.3.1 New Wireless Captioning System Debuts at Dallas

A proposta feita por West et al. (2009), apresentada na subseção anterior (*Wireless Captioning System* – 3.2.3), foi demonstrada em um cenário de larga escala em Janeiro de 2010 (Wallace e Toon 2010). O evento ocorreu no *Dallas Cowboys Stadium* na cidade de Arlington, estado do Texas nos Estados Unidos.

De acordo com o diretor de pesquisa em comunicações do *Georgia Institute of Technology*, Kirk J. Englehardt e o gerente da *Georgia Tech Research News & Publications*, John Toon, o Assistente de Comunicação foi adaptado aos sistemas já

existentes no estádio, onde, dessa forma, foi possível transmitir via *broadcast* o conteúdo acessível para cerca de 700 dispositivos móveis utilizando uma rede local sem fio (*wireless 802.11b*).

A Figura 5 ilustra o uso da solução proposta por West et al. (2009) durante o evento esportivo, onde os conteúdos acessíveis são obtidos através dos dispositivos móveis dos usuários.



Figura 5. Exibição das legendas durante um jogo de Futebol Americano no *Dallas Cowboys Stadium*.

O diretor do *Dallas Cowboys Stadium*, Scott Purcel, afirmou que ficou muito orgulhoso em oferecer um serviço que, utilizando um dispositivo móvel sem fio, fornece tecnologia assistiva personalizada para os seus clientes, incluindo aqueles que são surdos ou que possuam alguma deficiência auditiva.

De modo geral, esse capítulo apresentou alguns trabalhos com o propósito de minimizar a problemática enfrentada pelos deficientes auditivos em alguns cenários como, por exemplo, cinemas, ambientes acadêmicos e estádios de futebol. No entanto, nenhuma das soluções apresentadas viabiliza a geração automática e distribuição de janelas em línguas de sinais utilizando um avatar 3D nos cinemas, em especial, nas plataformas do Cinemas Digital. Portanto, a partir das limitações observadas e das dificuldades que os surdos encontram para participar desses ambientes, especificamente

dos cinemas, está sendo proposto neste trabalho uma solução tecnológica que permita o acompanhamento da tradução do conteúdo cinematográfico na sua língua natural de comunicação, ou seja, a LIBRAS. Essa solução será apresentada com detalhes no próximo capítulo.

4 Solução Proposta

Conforme apresentado nos Capítulos 1 e 3, existe na literatura um conjunto de trabalhos que endereçam os problemas de comunicação dos surdos. No entanto, algumas dessas propostas possuem custos operacionais altos (por exemplo, as soluções que utilizam intérpretes de LIBRAS) (Kato, et al. 2010) ou não atendem as necessidades reais desses usuários (por exemplo, as soluções baseadas em legendas ou em avatares 2D) (Tambassia, et al. 2012), (SONY 2012), (West, et al. 2009), não provendo soluções práticas ou viáveis para o problema.

Para tentar minimizar esses problemas, neste trabalho, é proposta e desenvolvida uma solução tecnológica, prática e eficiente, denominada CineLIBRAS, que tem como objetivo gerar automaticamente e distribuir vídeos de LIBRAS (sinalizados por um avatar 3D) em salas de cinema. A solução é composta por um conjunto de componentes de software. Cada componente constituinte desta solução possui uma responsabilidade bem definida. Dentre essas responsabilidades estão a extração de informações dos arquivos de legenda, aplicação das técnicas definidas por Araújo (2012) para traduzir textos da língua portuguesa para uma representação textual da LIBRAS, sincronização dos conteúdos multimídia e distribuição dos vídeos acessíveis dentro da sala de cinema.

Os detalhes dessa solução serão apresentados nas próximas seções. Na Seção 4.1 será apresentado o modelo conceitual do CineLIBRAS. Na Seção 4.2 será explicitado o protocolo de comunicação entre o CineLIBRAS e o sistema de distribuição de conteúdos cinematográficos (Fogo Player). Por fim, na Seção 4.3, será especificada a arquitetura dos componentes do CineLIBRAS.

4.1 Modelo Conceitual da Infraestrutura do CineLIBRAS

O modelo conceitual do CineLIBRAS é apresentado na Figura 6. De acordo com a ilustração, é possível observar que o cenário de estudo é composto basicamente por dois ambientes, um que controla os conteúdos multimídia, denominado ambiente de

controle, e o outro que representa o ambiente onde os usuários podem acompanhar a reprodução do filme, designado como ambiente de exibição.

O ambiente de controle, como o próprio nome indica, é o ambiente onde os conteúdos cinematográficos (vídeo, áudio e legendas) são gerenciados e coordenados de forma que essas mídias possam ser distribuídas e reproduzidas de forma adequada. Ao observar a Figura 6 percebe-se a existência de dois componentes principais, o Fogo Player e o CineLIBRAS.

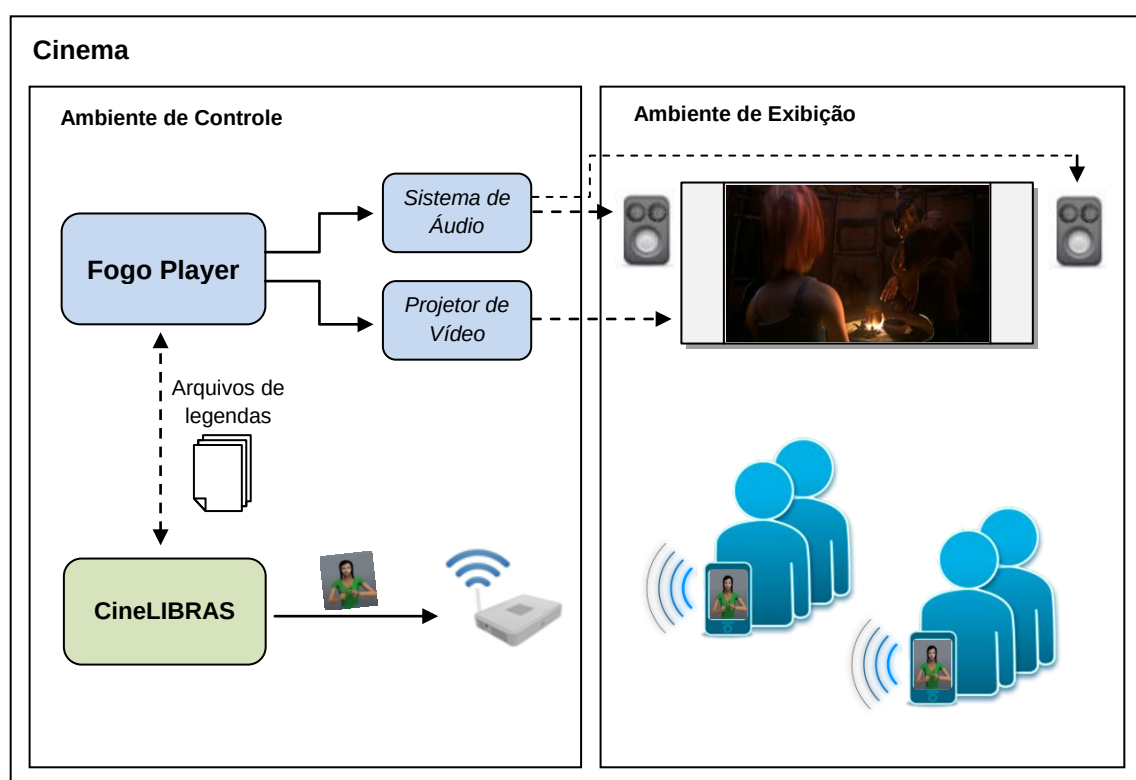


Figura 6. Modelo conceitual da infraestrutura do CineLIBRAS.

O Fogo Player, segundo Aquino Júnior et al. (2013), é uma solução que viabiliza a reprodução de vídeos com resolução 4K ($4096 \times 2160 \text{ pixels}$) – isto é, quatro vezes maior do que o *Ultra-High Definition* (UHD), com ou sem estereoscópio, de maneira sincronizada, mesmo quando diferentes partes do vídeos são transmitidas de fontes distintas. Essa solução é constituída por um conjunto de componentes de software que foram desenvolvidos para fornecer maior flexibilidade e mais qualidade na distribuição de conteúdos cinematográficos nos Cinemas Digitais (Aquino Júnior, et al. 2013).

Os principais componentes de software que constituem a arquitetura do Fogo Player foram divididos da seguinte forma: um conjunto de *Brick Players*¹, onde cada *brick player* é responsável pela recepção dos quadrantes de um vídeo, decodificação do conteúdo e reprodução dos *frames* decodificados; um *Controller*, que implementa alguns dos requisitos principais da coordenação do subsistema e garante a sincronização entre os *Brick Players*; o *Audio Player*, que controla o áudio do conteúdo sendo reproduzido; e, por fim, o *Fogo Streamer*, que é a implementação de um subsistema de *streamer* responsável pela transmissão de cada segmento do vídeo para um *Brick Player*. Maiores detalhes sobre o Fogo Player podem ser encontrados em (Aquino Júnior, et al. 2013).

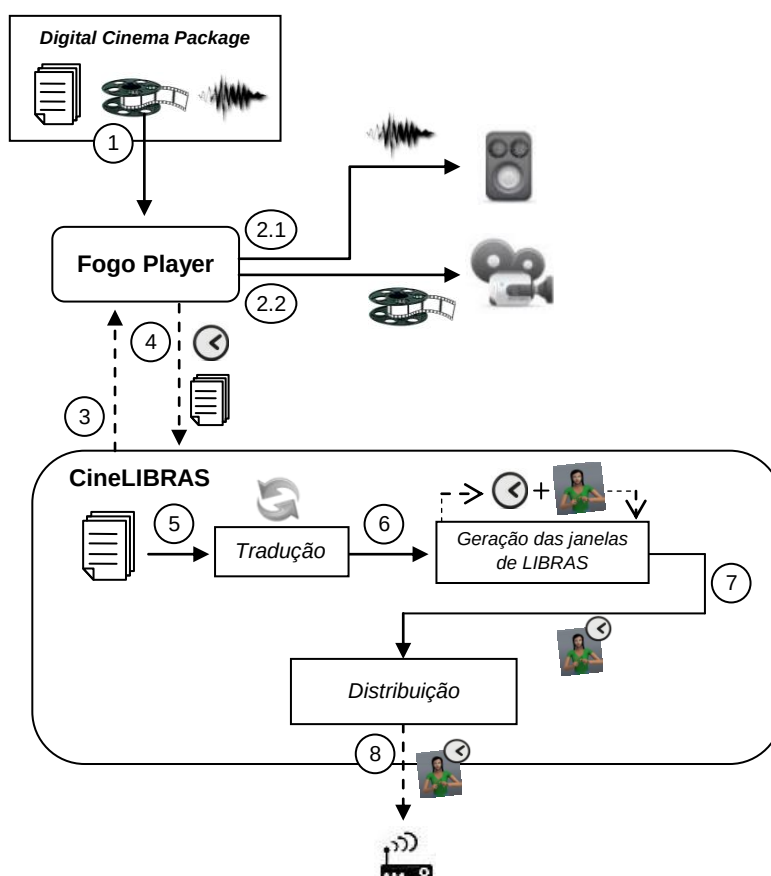


Figura 7. Sequência de execução para a geração e distribuição dos vídeos acessíveis.

¹ Segundo (Aquino Júnior, et al. 2013), um *brick* representa um quadrante de um vídeo. Dessa forma, um vídeo que contém quatro quadrantes, são necessários quatro *Brick Player* para receber, decodificar e exibir os frames de cada quadrante do vídeo.

O CineLIBRAS, por sua vez, é o módulo da solução proposta neste trabalho. O objetivo desse módulo é fornecer conteúdos acessíveis (isto é, janelas de LIBRAS) no ambiente de exibição conforme mostrado na Figura 6. Os conteúdos acessíveis são gerados a partir da trilha de legenda do filme onde, para obter essas legendas, o CineLIBRAS solicita ao módulo distribuidor de mídias cinematográficas (Fogo Player) por meio de um protocolo de comunicação (apresentado com detalhes na subseção 4.2) desenvolvido para este fim.

A Figura 7 apresenta alguns aspectos do fluxo de comunicação entre o Fogo Player e o CineLIBRAS. De acordo com a ilustração, percebe-se que algumas etapas do processo possuem uma determinada sequência de execução, isto é, uma hierarquia. No entanto, eventualmente, os passos 2 e 3 podem ocorrer de maneira invertida, pois, supondo que Fogo Player tenha recebido os pacotes DCP, mas que ainda não tenha dado início a reprodução dos conteúdos (passos 2.1 e 2.2), o CineLIBRAS, por sua vez, poderá fazer requisições ao Fogo Player para obter as trilhas de legenda, portanto, ocorrendo em momentos opostos.

Observando a sequência de estágios apresentada na Figura 7, para que os conteúdos acessíveis sejam transmitidos, de fato, é necessário passar por alguns estágios precedentes. A descrição de cada estágio pode ser compreendida da seguinte forma:

1. Os pacotes DCP contendo os fluxos de áudio, vídeo e legendas chegam no Fogo Player para serem processados;
2. O Fogo Player realiza os procedimentos de decodificação, sincronização e exibição desses conteúdos no ambiente de exibição (ver Figura 6);
3. O CineLIBRAS inicia o processo de comunicação com o Fogo Player, requisitando os arquivos de legenda e algumas informações de tempo (isto é, o relógio do Fogo Player), necessárias para a sincronização entre os módulos;
4. O Fogo Player envia os arquivos de legenda e as informações de tempo solicitadas pelo CineLIBRAS;
5. As legendas e suas respectivas etiquetas de tempo (isto é, *Presentation Timestamp* – PTS) são extraídas dos arquivos recebidos. Em seguida, essas

legendas são repassadas para o componente de tradução automática, que basicamente irá traduzir o texto do português para a LIBRAS, ou simplesmente *glosa*²;

6. A *glosa* é repassada para o componente de Geração das Janelas de LIBRAS, que irá utilizar as informações de tempo (relógio compartilhado do sistema e PTS das legendas) para sincronizar os vídeos de LIBRAS com o vídeo principal;
7. Os vídeos de LIBRAS são encaminhados ao componente de Distribuição, que irá transmitir esses vídeos para os usuários conectados no CineLIBRAS.

No ambiente de exibição, conforme visto na Figura 6, é o espaço onde os usuários (inclusive os deficientes auditivos) podem acompanhar a reprodução dos conteúdos cinematográficos. Esse ambiente é onde os usuários surdos poderão ter acesso aos vídeos de LIBRAS utilizando seus dispositivos móveis, desde que, esses dispositivos estejam de acordo com alguns requisitos especificados na Seção 4.3 (Arquitetura de Componentes do CineLIBRAS).

O modelo conceitual apresentado nessa seção mostrou de forma inteligível o cenário de atuação da solução proposta, assim como a organização dos principais módulos envolvidos. A próxima seção apresenta a especificação do protocolo de comunicação entre esses módulos, pois, conforme mencionado, essa comunicação é fundamental para que o CineLIBRAS possa obter os arquivos de legenda do filme e algumas informações de tempo para a sincronização dos conteúdos.

4.2 Protocolo de Comunicação com o Fogo Player

A comunicação entre os computadores acontece por meio da troca de mensagens. Essas mensagens são responsáveis por transportar dados e informações entre clientes e servidores. Conforme observado na Figura 6 (ambiente de controle), o CineLIBRAS utiliza o mecanismo de permutação de mensagens para obter algumas informações do Fogo Player como, por exemplo, os arquivos de legenda do filme e as referências do

² *Glosa* é uma representação de um texto na gramática de LIBRAS.

relógio. Na Figura 8, é apresentada a estrutura geral das mensagens utilizadas no processo de comunicação entre os módulos CineLIBRAS e FogoPlayer.

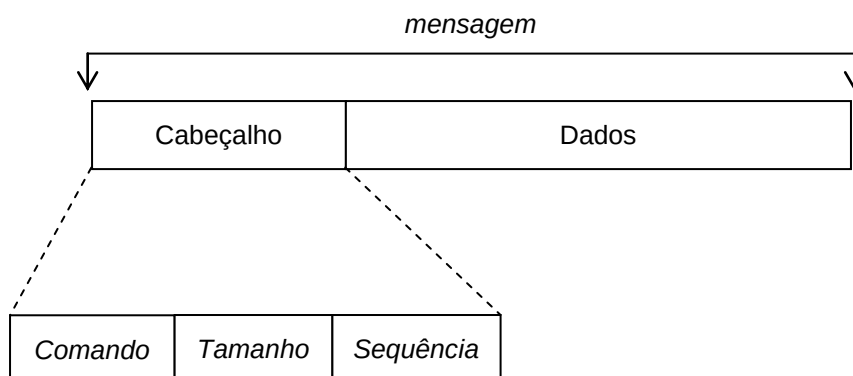


Figura 8. Estrutura da mensagem utilizada na comunicação com o Fogo Player.

De acordo com a Figura 8, as mensagens são divididas em duas partes, o Cabeçalho e os Dados. O Cabeçalho é utilizado para fornecer algumas informações sobre a mensagem. O campo Comando é utilizado para especificar o tipo de requisição que está sendo feita. O Tamanho deve ser usado para informar a extensão (em *bytes*) da carga útil da mensagem. No campo Sequência, deve ser informada a quantidade de mensagens restantes para transmitir as informações solicitadas. Esse parâmetro é fundamental para enviar, por exemplo, os metadados dos arquivos de legenda, pois, dependendo do tamanho e da quantidade de arquivos, são necessárias mais de uma mensagem. No outro segmento da mensagem, isto é, Dados, são transportadas as informações inerentes ao contexto da comunicação como, por exemplo, o PCR e os metadados dos arquivos (XML) de legenda.

Para atender uma requisição a partir de uma mensagem, é necessário interpretá-la, isto é, o receptor deve compreendê-la através dos parâmetros contidos no cabeçalho e no conteúdo da carga útil. Portanto, para interpretar uma mensagem é preciso estabelecer um conjunto de critérios bem definidos, baseando-se na coerência e na consistência das informações. Dessa forma, com a finalidade de atender aos requisitos da comunicação entre o CineLIBRAS e o Fogo Player, foi desenvolvido um Protocolo de Comunicação (Araújo, Pedroza, & Mesquita, 2003).

Segundo Araújo et al. (2003), um Protocolo de Comunicação é definido como um conjunto de regras que governa o formato e o significado de quadros, mensagens ou pacotes que são trocados através de entidades pares. Dessa forma, a especificação do Protocolo de Comunicação entre o CineLIBRAS e o Fogo Player foi elaborada com base nas propriedades de comunicação definidas por Araújo et al. (2003).

De acordo com Araújo et al. (2003), as principais propriedades de comunicação entre dois pares são: segurança e vivacidade. A primeira trata dos aspectos sintáticos das funcionalidades definidas, garantindo a coerência da comunicação e a estabilidade do protocolo. Na segunda, são abordadas algumas características no que se refere a correção semântica, assegurando a disponibilidade do protocolo. A Tabela 2 apresenta a especificação das mensagens utilizadas no Protocolo de Comunicação entre o CineLIBRAS e o Fogo Player.

Tabela 2. Especificação das mensagens do protocolo de comunicação com o Fogo Player.

<i>REQUEST</i>		<i>RESPONSE</i>	
Comando	Descrição	Comando	Descrição
IDENTITY	Solicita uma autenticação no servidor por meio do <i>Unique Identification</i> (UID) da aplicação	IDENTITY	Se a autenticação for realizada com sucesso, retorna OK, caso contrário, responde ERRO
CLOCK	Solicita a hora atual do servidor para ajustar o sincronismo dos conteúdos acessíveis	PCR	Envia o <i>Program Clock Reference</i> (PCR) do módulo servidor
SUBTITLE	Solicita os arquivos de legenda do filme	FILE	Envia os metadados dos arquivos de legenda
		INITIALIZE	Indica que a reprodução do filme foi iniciada
		FINALIZE	Informa o encerramento da reprodução do filme

Observando a Tabela 2, é possível observar que as mensagens estão agrupadas em duas categorias: *REQUEST* e *RESPONSE*. Cada mensagem do grupo *REQUEST*

está relacionada com uma mensagem *RESPONSE*, com exceção das mensagens *INITIALIZE* e *FINALIZE*, que são utilizadas pelo Fogo Player para informar o início e o fim da reprodução de um filme. Uma vez que o início e fim da reprodução não depende do CineLIBRAS, não há necessidade de requisições para essas mensagens. Para ilustrar o funcionamento dessas mensagens, na Figura 9, é ilustrado um exemplo de troca de mensagens entre o CineLIBRAS e o Fogo Player com base no protocolo proposto.

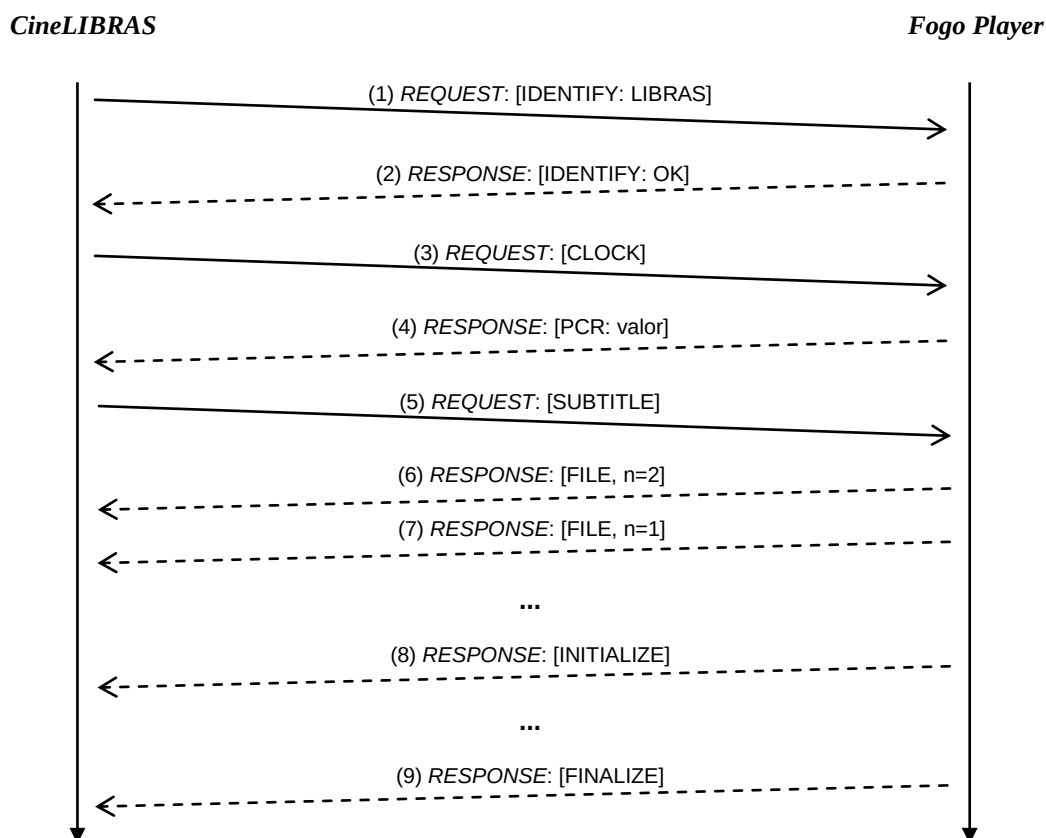


Figura 9. Sequência de mensagens trocadas entre o CineLIBRAS e o Fogo Player.

De acordo com a Figura 9, inicialmente, o CineLIBRAS solicita sua autenticação no Fogo Player por meio de uma mensagem *REQUEST* com o comando *IDENTITY* e seu *Unique Identification* (UID). Em seguida, o Fogo Player envia uma mensagem *RESPONSE* informando que a autenticação foi bem sucedida. Dessa forma, após receber a confirmação, o CineLIBRAS está habilitado para fazer novas requisições e receber os conteúdos enviados pelo Fogo Player.

Posteriormente, na mensagem 3, o CineLIBRAS solicita as informações referentes ao relógio do Fogo Player que, como resposta, envia uma mensagem contendo seu *Program Clock Reference* (PCR³). Em seguida, uma mensagem SUBTITLE é enviada pelo CineLIBRAS para solicitar os arquivos (XML) de legenda do filme (ver Seção 4.3). A resposta dessa requisição são duas mensagens enviadas uma após a outra, indicando que o arquivo de legenda teve que ser dividido para que pudesse ser transmitido.

Após receber o arquivo de legenda, o CineLIBRAS fica aguardando uma mensagem com o comando INITIALIZE informando que a apresentação do filme foi iniciada. Quando essa mensagem é recebida, o CineLIBRAS inicia o procedimento de geração e distribuição dos conteúdos acessíveis na sala do cinema. Da mesma forma, quando o filme for encerrado, o Fogo Player dispara uma mensagem com o comando FINALIZE ao CineLIBRAS que, ao receber essa mensagem, finaliza todos os seus processos.

O Protocolo de Comunicação desenvolvido nesta etapa do trabalho é fundamental para sincronizar os componentes de apresentação e geração de conteúdos em LIBRAS (Fogo Player e CineLIBRAS, respectivamente). No entanto, como esses componentes estão em fase de desenvolvimento, existem alguns aspectos desse protocolo que estão sendo avaliados e aperfeiçoados. Essas melhorias devem ser incorporadas nas próximas versões do protocolo.

4.3 Arquitetura de Componentes do CineLIBRAS

O CineLIBRAS é composto por um conjunto de componentes de *software* que tem como objetivo gerar conteúdos acessíveis (janelas de LIBRAS) para usuários surdos. em salas de cinema Esses conteúdos produzidos pela solução são distribuídos para os usuários que podem acompanhar a tradução de um filme em LIBRAS a partir do seu dispositivo móvel.

³ Informação de tempo utilizada na compressão de vídeo digital para indicar o *System Time Clock's* (STC) ao decodificador.

Os requisitos básicos para que os usuários recebam os vídeos de LIBRAS em seus dispositivos é que eles estejam conectados ao CineLIBRAS através de uma rede local e que possuam players de vídeo instalados com suporte a recepção de fluxos MPEG-2 TS via *streaming* HTTP (*HyperText Transfer Protocol*). A Figura 10 apresenta a arquitetura do CineLIBRAS.

De acordo com a Figura 10, a solução funciona da seguinte forma. Inicialmente, o CineLIBRAS estabelece um processo de comunicação com o FogoPlayer e requisita os arquivos de legenda do filme. Os arquivos de legendas são então encaminhados para o componente *Extrator de Legendas* que extrai o texto das legendas e as informações de sincronização de cada sentença do texto. As sentenças são então encaminhadas para o componente *Tradutor*, um subcomponente do *Gerador de LIBRAS*, que, traduz o texto em língua portuguesa para uma representação textual na gramática de LIBRAS, denominada *glosa*.

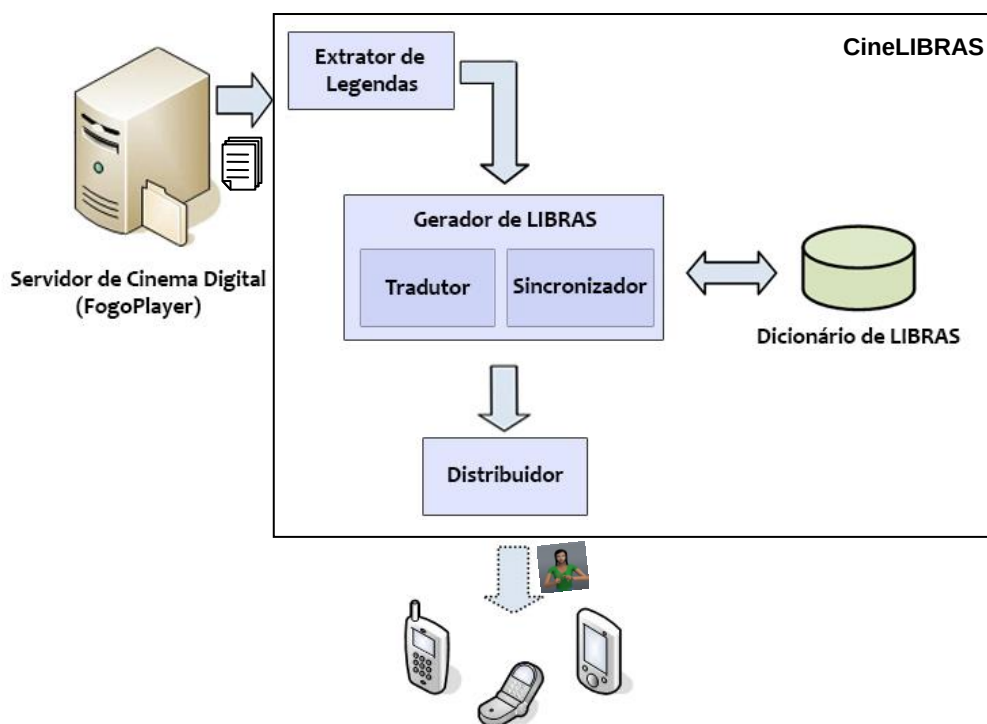


Figura 10. Arquitetura do CineLIBRAS.

Após o processo de tradução, o componente *Gerador de LIBRAS* converte a sequência de glosas em um vídeo de LIBRAS, com o auxílio do *Dicionário de LIBRAS*, e sincroniza esse vídeo com o conteúdo do filme usando as informações de sincronização extraídas do arquivo de legenda. Por fim, o vídeo de LIBRAS é encaminhado para o componente de *Distribuidor* que pelo transmite o vídeo acessível para os dispositivos dos usuários conectados.

Nas próximas subseções, os componentes do CineLIBRAS serão apresentados em maiores detalhes.

4.3.1 Extrator de Legendas

O componente denominado Extrator de Legendas é responsável por extrair o texto (na linguagem oral/escrita) e informações de sincronização de cada sentença (fundamentais para a sincronização destas com o vídeo principal) dos arquivos de legenda definidos pela *Digital Cinema Initiatives* (DCI). Após extrair o texto e as informações de sincronização das legendas, o Extrator de Legendas repassa essas informações para o componente Gerador de LIBRAS.

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <DCSubtitle Version="1.0">
3   <SubtitleID>EAE5652F-0EE5-4009-ACAB-EAE5B450C357</SubtitleID>
4   <MovieTitle>Sintel</MovieTitle>
5   <ReelNumber>1</ReelNumber>
6   <Language>Portuguese</Language>
7   <LoadFont Id="Arial" URI="arial.ttf" />
8   <Font Id="Arial" Effect="shadow" Size="42">
9     <Subtitle SpotNumber="001" TimeIn="00:01:47:250" TimeOut="00:01:50:500">
10       <Text VAlign="bottom" VPosition="15.00">
11         Esta lâmina tem um passado negro.
12       </Text>
13     </Subtitle>
14     <Subtitle SpotNumber="002" TimeIn="00:01:51:800" TimeOut="00:01:55:800">
15       <Text VAlign="bottom" VPosition="9.00">
16         Derramou demasiado sangue inocente.
17       </Text>
18     </Subtitle>
19   </Font>
20 </DCSubtitle>
```

Figura 11. Estrutura do arquivo de legendas recomendado pela DCI.

Os arquivos de legenda processados pelo Extrator de Legendas são compatíveis com a especificação do *Digital Cinema System Specification* (DCSS), que são representados usando o formato XML (*eXtensible Markup Language*) recomendado pela *World Wide Web Consortium* (W3C). De acordo com a DCI, um pacote *Digital Cinema Package* (DCP) pode conter um ou mais arquivos de legenda. A Figura 11 apresenta um exemplo de arquivo de legenda de acordo com a especificação do DCI.

A partir da linha 9 do arquivo, dentro do campo *Subtitle*, são apresentados os parâmetros que descrevem o conteúdo (texto) da legenda e as suas marcações de tempo. Os parâmetros *TimeIn* e *TimeOut* representam o instante de tempo em que a legenda deve ser exibida na tela e retirada da tela, respectivamente, no formato *HH:MM:SS:Centésimos*, enquanto que o parâmetro *Text* representa a sentença que deve ser exibida na tela.

4.3.2 Gerador de LIBRAS

Conforme apresentado na Figura 8, o componente *Gerador de LIBRAS* é composto por dois subcomponentes, o *Tradutor* e o *Sincronizador*. O Tradutor tem como finalidade converter as sentenças textuais recebidas do *Extrator de Legendas* para glosa, e o Sincronizador é responsável por converter uma sentença representada em glosa para um vídeo em LIBRAS com o auxílio do *Dicionário de LIBRAS*.

Nas próximas subseções, os componentes de tradução e sincronização são apresentados com mais detalhes.

4.3.2.1 Tradutor

Conforme mencionado anteriormente, as línguas orais (baseada em sons) possuem um conjunto de regras gramaticais que diferem completamente das línguas de sinais (constituída por movimentos gestuais e expressões faciais). Portanto, como o texto da legenda é baseado na linguagem oral (isto é, língua portuguesa), é necessário que o mesmo seja traduzido para a língua de sinais (nesse caso, LIBRAS).

O processo de tradução de uma representação textual da linguagem oral para a língua de sinais é realizado pelo componente denominado Tradutor. Para realizar esta tarefa, foi utilizada a implementação do Tradutor desenvolvida por Araújo (2012). Segundo Araújo (2012), esse procedimento de tradução automática combina métodos de compressão estatística utilizados para classificar os *tokens* (palavras) de entrada, estratégias de simplificação textual para reduzir a complexidade do texto de entrada e um conjunto de regras morfológicas e sintáticas. Maiores detalhes sobre o processo de tradução automática podem ser encontrados em Araújo (2012).

Por fim, após realizar os procedimentos necessários para a tradução, o Tradutor produz uma saída que consiste na representação em *glosa* do texto de entrada, por exemplo, dada a sentença "eu vou para a praia" escrita na língua portuguesa, o Tradutor gera a saída "EU PRAIA IR" em LIBRAS. Dessa forma, o resultado gerado pelo componente de tradução é repassado para o *Sincronizador*, que é o componente responsável pela conversão de uma sequência de *glosas* em vídeos de LIBRAS. O Sincronizador será apresentado com detalhes na próxima subseção .

4.3.2.2 Sincronizador

Conforme mencionado anteriormente, o Sincronizador tem como função converter a sequência de glosas em um vídeo de LIBRAS mantendo as relações de sincronismo com as mídias do filme. Para realizar essa tarefa, ele utiliza um Dicionário de LIBRAS que contém uma representação visual (por exemplo, uma animação ou um arquivo de vídeo) para cada sinal em LIBRAS. Mais especificamente, o componente recebe a sequência de glosas e realiza uma consulta no Dicionário de LIBRAS, recuperando um vídeo para cada glosa (sinal) da sentença correspondente. Após a recuperação dos vídeos de LIBRAS, esses vídeos são combinados (concatenados), gerando uma trilha de vídeo em LIBRAS. O sincronismo entre o fluxo de legenda e a trilha de LIBRAS é mantido, adicionando etiquetas de tempo (pontos de sincronização) aos vídeos de LIBRAS similares as etiquetas de tempo utilizadas na legenda.

Dessa forma, a estratégia de sincronização utilizada pela solução é baseada no modelo de sincronização baseado no eixo do tempo (Blakowski e Steinmetz 1996),

onde o relógio global do filme é utilizado como referência para gerar as etiquetas de tempo de apresentação (PTS) de todas as mídias do filme: vídeo principal, áudio, trilhas de legendas e de LIBRAS.

Para que as transições entre os vídeos de LIBRAS (consecutivos) sejam suaves, uma configuração neutra foi definida no início e no final de cada sinal e durante os intervalos de silêncio. Essa configuração foi definida de acordo com a sugestão de especialistas em LIBRAS, e consiste em posicionar o avatar 3D com os braços estendidos para baixo e com uma expressão facial neutra.

Além disso, caso não exista um vídeo para algum dos sinais da sentença no Dicionário de LIBRAS, um vídeo é gerado para aquele sinal a partir da soletração da sua glosa. Essa estratégia é utilizada para evitar lacunas na representação das sentenças em LIBRAS e é a mesma estratégia utilizada pelos surdos e pelos intérpretes de LIBRAS para representar palavras ou termos que não possuem sinais próprios, como, por exemplo, os nomes próprios e os termos técnicos. Por exemplo, em LIBRAS, o termo "CINELIBRAS" é representado como "C-I-N-E-L-I-B-R-A-S".

Após a geração da trilha (vídeo) de LIBRAS, o *Sincronizador* repassa essa trilha para o componente *Distribuidor* que transmite essa trilha para os clientes conectados no CineLIBRAS.

4.3.3 Dicionário de LIBRAS

O *Dicionário de LIBRAS* é uma base de dados que contém um conjunto de vídeos (no formato MPEG2-TS) representados por um avatar 3D em LIBRAS. Essa é uma característica fundamental para a geração do conteúdo acessível, pois, a partir deste dicionário, outros componentes que necessitem obter a representação em LIBRAS de uma determinada sentença, podem acessá-lo para adquirir o sinal correspondente.

No Dicionário de LIBRAS, os sinais são representados por um arquivo de vídeo, e cada um deles possui uma chave associada. Dessa forma, o dicionário pode ser definido como um conjunto de tuplas t no seguinte formato:

$$t = \langle g, v \rangle,$$

Onde:

- *g* é a glosa (ou chave) do sinal (i.e., a representação textual do sinal);
- *v* é a representação visual do sinal;

Como cada sinal possui uma chave distinta, torna-se relativamente simples customizá-los a qualquer momento, sem que ocorram mudanças no procedimento de acesso aos sinais. Além disso, uma característica importante de usar o esquema de tuplas é que, como algumas regiões do país possuem diferentes formas para representar alguma coisa (isto é, regionalismo), a estrutura do Dicionário de LIBRAS consegue se adaptar facilmente.

4.3.4 Distribuidor

O componente denominado *Distribuidor* é responsável pela transmissão do vídeo (trilha) gerado pela solução nas salas de cinema. Esse processo de transmissão é realizado utilizando uma rede local, onde, para receber o conteúdo gerado, os clientes precisam fazer uma solicitação de recebimento informando seu endereço IP (*Internet Protocol*) na rede. Além disso, para que o dispositivo seja capaz de exibir o conteúdo, é necessário que este tenha suporte para receber, decodificar e apresentar fluxos de vídeo via *streaming* UDP.

Durante o processo de transmissão do conteúdo gerado, o Distribuidor utiliza uma abordagem baseada em refletores. Essa abordagem foi utilizada para tentar manter uma boa sincronia entre os clientes conectados na solução. Dessa forma, o componente de distribuição utiliza a mesma fonte de dados e, replica (reflete) os pacotes de vídeo dessa fonte para cada usuário conectado. Os dados são enviados, ciclicamente, em unidades de 7 (sete) pacotes TS (*Transport Stream*) de 188 *bytes* cada (isto é, 1.316 *bytes* no total) por vez, para cada cliente. Essa estratégia de distribuição tende a reduzir o retardo de transmissão entre os clientes, uma vez que poucos bytes são transmitidos, por vez, para cada cliente, além de reduzir a fragmentação dos datagramas na rede, uma vez que tamanho do pacote, 1316 *bytes* ($=7 \times 188$ *bytes*), aproxima-se do MTU

(*Maximum Transmission Unit* – unidade máxima de transmissão) da maioria das redes baseadas em pacotes, que, normalmente, definem um MTU de 1500 bytes.

Essa abordagem é fundamental em ambientes públicos como o cinema, onde a exibição dos conteúdos multimídia é compartilhada por usuários surdos e ouvintes. Portanto, essa estratégia de distribuição dos vídeos de LIBRAS em dispositivos de segunda tela é interessante, pois o vídeo de LIBRAS é disponibilizado apenas para o *display* do usuário surdo, tornando a sessão do cinema inclusiva.

5 Resultados

Este capítulo apresenta alguns resultados obtidos após a realização de experimentos com um protótipo da solução proposta. O objetivo foi investigar o nível de inteligibilidade⁴ dos conteúdos gerados pela solução e o grau de satisfação dos surdos em utilizar um dispositivo de segunda tela para acompanhar a tradução (em LIBRAS) do conteúdo do filme.

Para isso, será apresentado um conjunto de testes realizados com usuários surdos, confrontando a solução proposta com outras três abordagens: (1) utilizando legendas no formato texto, (2) utilizando uma janela de LIBRAS gerada por um intérprete de LIBRAS e sobreposta a tela do filme, e (3) utilizando uma janela de LIBRAS gerada pelo avatar 3D e também sobreposta a tela do filme, de forma similar a abordagem (2).

5.1 Teste de Inteligibilidade dos Conteúdos Gerados

Durante algum tempo, a maior parte dos esforços na construção de soluções tecnológicas se concentravam em melhorar a qualidade e aumentar o desempenho dos sistemas na perspectiva computacional. No entanto, mesmo dispondo de numerosos recursos, muitas soluções obtinham resultados insatisfatórios do ponto de vista do seu público alvo. Em consequência disso, novas técnicas de avaliação foram criadas para ajudar no processo de construção dessas soluções. Uma delas é o desenvolvimento de um protótipo da solução e a realização de experimentos envolvendo de forma mais efetiva a participação do usuário final (Steves, et al. 2001).

Em virtude disso, foi realizado um experimento com usuários surdos. Esse experimento aconteceu em duas etapas. Na primeira etapa, um conjunto de usuários surdos avaliou o uso de legendas no formato texto e de um avatar 3D posicionado na tela, e na segunda etapa, um outro conjunto de usuários avaliaram o uso de janelas com

⁴ Qualidade, caráter de uma coisa inteligível (Dicio 2013).

intérprete de LIBRAS e a solução proposta (CineLIBRAS). O filme utilizado nas duas etapas do experimento para os dois tratamentos foi o Sintel (2010), um filme de animação que possui classificação livre e duração de aproximadamente quinze minutos. A Figura 12 mostra uma visão esquemática de como o experimento foi estruturado.

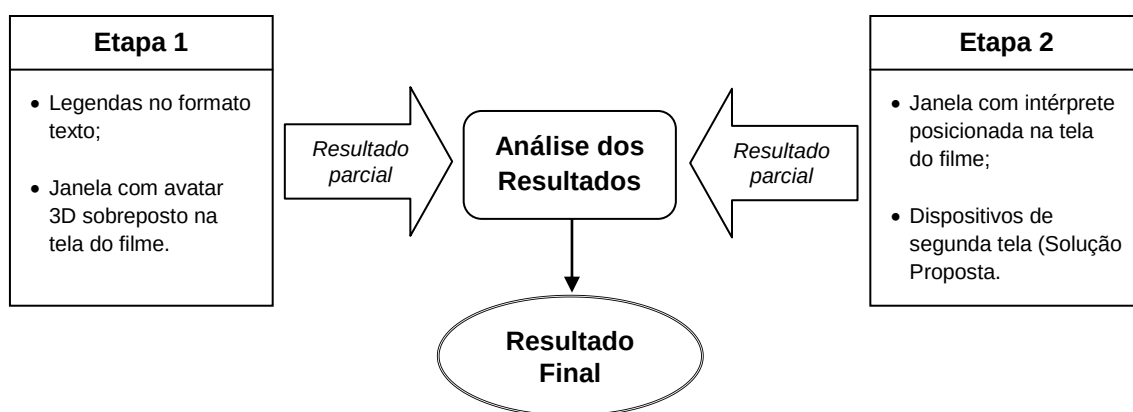


Figura 12. Modelo conceitual do experimento realizado.

As seções a seguir apresentam uma descrição detalhada de cada etapa desse experimento. Na primeira parte, é apresentado o processo de avaliação conduzido na "Etapa 1" (ver Figura 12), enquanto que na segunda parte, é descrito o procedimento do teste realizado na "Etapa 2". Por fim, na terceira parte, é apresentada uma análise comparativa descritiva dos resultados gerados nas duas etapas do experimento.

5.1.1 Avaliação da primeira etapa do experimento

A primeira etapa do experimento foi realizada nos dias 29 de Julho e 02 de Agosto de 2012, na Fundação Centro Integrado de Apoio a Pessoa com Deficiência (FUNAD), localizada na cidade de João Pessoa - PB. Esta etapa contou com a participação de vinte usuários surdos do curso de formação em LIBRAS da FUNAD. Esses usuários foram divididos aleatoriamente em dois grupos de dez usuários, onde um grupo deveria assistir o filme utilizando o recurso da legenda no formato texto, e o outro, visualizando a tradução do conteúdo através de uma janela de LIBRAS (com um avatar 3D) sobreposta na tela do filme.

Inicialmente, os usuários surdos foram preparados para participar do experimento. Na fase de preparação, o propósito do teste foi apresentado aos usuários e, em seguida, eles receberam um questionário (apresentado no Anexo A) e foram convidados a preencher a primeira parte com suas informações pessoais, contando com o auxílio de intérpretes de LIBRAS, .

O grupo de usuários que participou desta etapa era heterogêneo, composto por onze mulheres e nove homens. A faixa etária desse grupo foi de 13 a 56 anos, com média de 28,6 anos de idade. O grau de escolaridade dos usuários também foi observado e é apresentado na Figura 13.

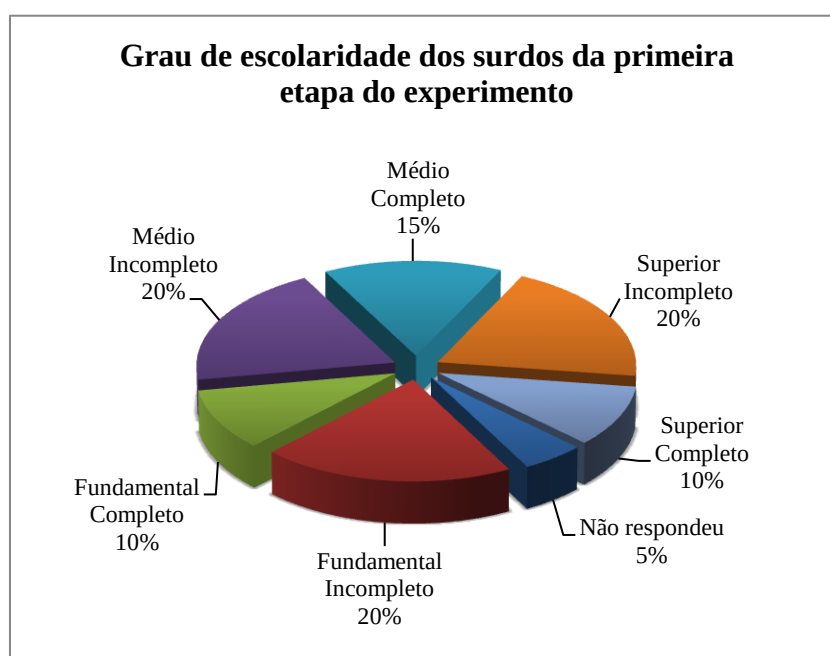


Figura 13. Grau de escolaridade dos vinte usuários surdos que participaram da primeira etapa do experimento.

Logo após o preenchimento da primeira parte do questionário, os usuários surdos foram convidados para assistir o filme utilizado no experimento. Após a exibição do filme, os usuários de cada grupo foram convocados a responder a segunda parte do questionário, que estava relacionada ao conteúdo apresentado. A Figura 14 mostra alguns momentos durante esta etapa do experimento.



Figura 14. Momentos durante a realização da primeira etapa do experimento: (a) Alguns usuários que participaram do teste, (b) Fase de preparação dos usuários, (c) Durante a avaliação do conteúdo com o recurso das legendas textuais, e (d) Alguns usuários surdos assistindo ao conteúdo com o avatar posicionado na tela.

Os resultados obtidos na segunda parte do questionário são apresentados na Tabela 3. Conforme pode ser observado nessa tabela, é possível perceber que os surdos que assistiram o filme utilizando o recurso das legendas tiveram um nível de entendimento inferior com relação aos usuários que utilizaram o avatar 3D posicionado na tela, isto é, com o recurso das legendas textuais, os surdos tiveram um percentual de acerto de 35%, enquanto os surdos que acompanharam a tradução através do avatar 3D conseguiram um percentual de 83% de acerto.

Tabela 3. Nível de acerto das perguntas sobre o conteúdo na primeira etapa do experimento.

	Média	Desvio Padrão	Média (%)
Legendas	1,40	0,24	35%
Avatar na Tela	3,30	0,21	83%

Outro aspecto observado nessa etapa do experimento foi o nível de satisfação dos sinais representados pelo avatar 3D. Nessa abordagem, os surdos tiveram que avaliar dois pontos essenciais: a facilidade de compreensão dos sinais realizados e a aderência à gramática de LIBRAS. Dessa forma, os surdos deveriam escolher um número numa escala entre 1 e 6, onde 1 significa que os sinais foram mal representados e 6, que foi muito fácil compreender a sinalização. O resultado obtido nesse aspecto é apresentado na Tabela 4.

Tabela 4. Classificação do nível de compreensão dos sinais realizados pelo avatar 3D.

	Quesito	Média	Desvio Padrão
Avatar na Tela	Compreensão	4,60	0,30
	Gramática	4,60	0,27

Analisando a tabela acima, é possível perceber que os sinais realizados pelo avatar 3D foram bem representados, isto é, os sinais executados pelo avatar obtiveram um valor médio de 4,6. Em relação a aderência gramatical dos sinais em LIBRAS feitos pelo avatar, a média atribuída pelos surdos foi de 4,60, com desvio padrão de 0,27. Isso provavelmente significa que os conteúdos gerados estão razoavelmente aderentes com a gramática de LIBRAS.

5.1.2 Avaliação da segunda etapa do experimento

A segunda etapa do experimento, que inclui a avaliação da solução proposta, aconteceu nos dias 04 e 05 de Setembro de 2013 no Instituto Nacional de Educação de Surdos (INES), localizado na cidade do Rio de Janeiro - RJ. Nesta etapa, houve a participação de vinte e quatro alunos surdos do sistema de ensino bilíngue do INES. Esses surdos foram divididos em dois grupos distintos para avaliar as duas abordagens usadas nesta etapa. Na primeira, os surdos tiveram que utilizar a solução proposta neste trabalho, isto é, visualizar a tradução do conteúdo utilizando um dispositivo de segunda tela. Na segunda, os surdos precisaram acompanhar a tradução do conteúdo em uma janela com intérprete de LIBRAS sobreposta a tela de exibição do filme.

O grupo de surdos que participou desta etapa foi constituído por dezenove mulheres e cinco homens (totalizando 24 surdos). No entanto, para manter compatibilidade com a quantidade de amostras coletadas na primeira etapa do experimento, após a organização dos dados coletados nesta etapa, quatro amostras foram removidas ao acaso, isto é, com probabilidade igualmente distribuída elas. Dessa forma, apenas vinte amostras foram contabilizadas.

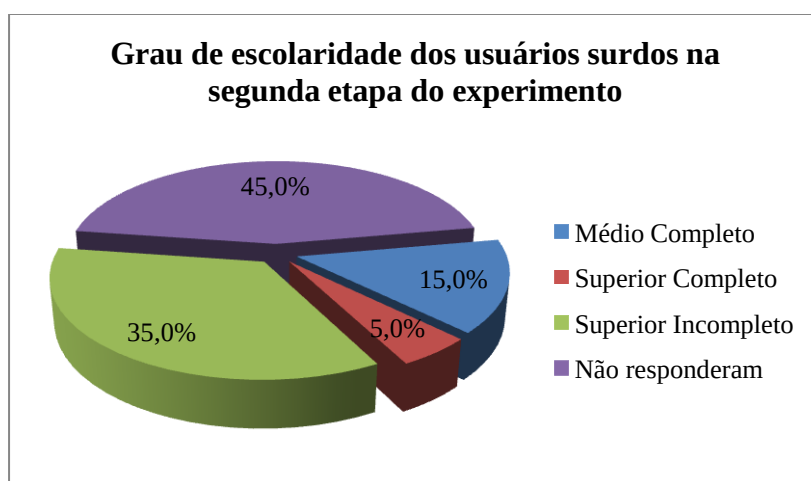


Figura 15. Nível de escolaridade dos surdos que participaram da segunda etapa do experimento.

Após reformular o conjunto das amostras, o grupo resultante passou a ser formado por quinze mulheres e cinco homens. A faixa etária desses usuários foi de 18 a

57 anos, com média correspondente a 32,75 anos, e um desvio padrão de 12,13. O grau de escolaridade dos participantes também foi analisado e é apresentado na Figura 15.

A primeira abordagem contou com a participação de onze surdos (sendo contabilizados apenas dez em virtude da compatibilidade entre as amostras). Cada surdo participante teve que utilizar um dispositivo móvel com capacidade para receber as janelas de LIBRAS, isto é, um *software* instalado no dispositivo capaz de receber (por meio do protocolo UDP), decodificar e exibir fluxos de vídeo. Nos testes, foi usado o VLC (2013) para esta função. Os dispositivos móveis utilizados foram da marca Apple Inc. (2013), referentes ao modelo iPad 2.

Na segunda abordagem desta etapa, houve a colaboração de treze usuários surdos (considerando, da mesma forma, apenas dez amostras nas análises). Para esse grupo de surdos, a ideia foi disponibilizar a tradução do conteúdo em uma janela com um intérprete de LIBRAS, sendo essa janela posicionada na mesma tela de exibição do filme. As diretrizes para o posicionamento dessa janela está definido na Associação Brasileira de Normas Técnicas (ABNT 2005), que determina que a altura da janela deve ser no mínimo metade da altura da tela, e a largura deve ocupar no mínimo a quarta parte da largura da tela.

Em um primeiro momento, antes da exibição do filme, foram apresentados os objetivos do experimento. Em seguida, foi entregue a cada participante uma cópia do questionário, que é constituído por três seções: a primeira foi referente a alguns dados pessoais; a segunda, sobre algumas questões referentes ao conteúdo do filme, e a terceira, sobre o nível de compreensão dos sinais realizados pelo avatar 3D e a aderência à gramática de LIBRAS.

Para ajudar os surdos no preenchimento dos questionários, foi necessário convidar alguns intérpretes de LIBRAS do próprio INES. Dessa forma, os usuários surdos foram instruídos a responder apenas a primeira parte do questionário antes de assistir o conteúdo. Logo em seguida, após o preenchimento dessa parte inicial, foi comunicado que seria dado início a reprodução do filme, e que eles deveriam acompanhar a tradução utilizando os dispositivos móveis. A Figura 16 apresenta alguns momento durante essa etapa do experimento.

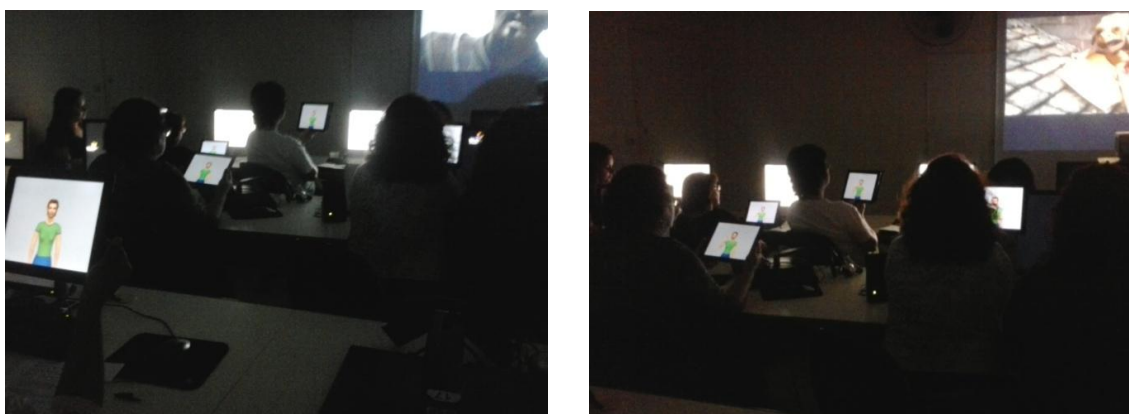


Figura 16. Momentos durante o teste com a solução proposta.

Após a exibição do filme, nas duas abordagens investigadas, os usuários foram convidados a responder as perguntas referentes a segunda parte do questionário. O objetivo dessa parte foi investigar o grau de compreensão do conteúdo exibido no filme. Dessa forma, em uma das seções, os surdos tiveram que responder a quatro perguntas, sendo todas de múltipla escolha e, estritamente relacionadas a algo, ou algum evento ocorrido no filme. O resultado dessa avaliação é apresentados na Tabela 5.

Tabela 5. Nível de acerto das perguntas sobre o conteúdo na segunda etapa do experimento.

	Média	Desvio Padrão	Média (%)
Solução Proposta	2,90	1,20	73%
Intérprete na Tela	3,20	0,79	80%

Analisando os dados apresentados na tabela acima, é possível perceber que os surdos que utilizaram a abordagem da janela com intérprete de LIBRAS posicionada na tela, tiveram um nível de acerto um pouco superior aos resultados obtidos pela solução proposta. Isto é, enquanto utilizando a solução proposta, a média de acerto das perguntas foi de 2,90, com desvio padrão de 1,20. Já com a janela do intérprete fixada na tela, a média de acertos obteve uma média de 3,20 e desvio padrão de 0,79. Em

termos percentuais, a abordagem com o intérprete na tela atingiu 80% de aproveitamento, enquanto com a solução proposta esse índice foi de 73%.

Uma característica investigada na segunda parte do questionário, apenas com a solução proposta, foi o nível de satisfação dos surdos em relação do avatar 3D. Nesse quesito, os surdos deveriam classificar a facilidade de compreensão dos sinais realizados pelo avatar e a aderência à gramática de LIBRAS. O método utilizado para classificação foi baseado numa escala numérica, onde eles deveriam atribuir um número entre 1 e 6, onde 1 significa muito ruim e 6 excelente, para representar sua resposta. O resultado dessa investigação é exposto na Tabela 6.

Tabela 6. Classificação do nível de compreensão dos sinais realizados pelo avatar 3D e a aderência a gramática de LIBRAS utilizando a solução proposta.

	Quesito	Média	Desvio Padrão
Solução Proposta	Compreensão	3,40	1,71
	Gramática	2,40	1,35

Examinando os dados apresentados na tabela acima, percebe-se que, em relação a gramática de LIBRAS, o avatar não atingiu um resultado satisfatório, ou seja, obtendo um valor médio de 2,4 com o desvio padrão de 1,35. No que se refere a compreensão dos sinais feitos pelo avatar, o resultado foi um pouco melhor, resultando num valor médio de 3,40 e desvio padrão de 1,71.

O último ponto investigado nesta etapa do experimento, foi a respeito do nível de satisfação do surdo em utilizar um dispositivo móvel para acompanhar a tradução do conteúdo. Nessa investigação, os usuários deveriam responder duas perguntas: a primeira sobre o nível de complexidade em utilizar um dispositivo móvel, e a segunda, quanto ao grau de dificuldade para acompanhar, simultaneamente, o filme e a tradução em LIBRAS nos dispositivos móveis. A resposta de cada item deveria ser um número entre 1 e 6 (sendo 1 muito difícil e 6 muito fácil) que representasse seu nível de satisfação. Os resultados obtidos nesse ponto são apresentados na Tabela 7.

Fazendo uma análise nos resultados apresentados na tabela abaixo, percebe-se que a média sobre a facilidade em utilizar um dispositivo móvel foi de 3,80, com um desvio padrão de 1,55. Um ponto importante sobre esse resultado é que os alunos surdos do INES que participaram desse teste, já possuíam um dispositivo móvel do mesmo modelo do especificado nesse teste, portanto, isso pode justificar uma média proeminente.

Em relação a segunda pergunta, o método utilizado na solução proposta, isto é, usar um dispositivo móvel para apresentar a tradução do conteúdo em LIBRAS, obteve uma média de 2,90, e um desvio padrão de 1,97. Após responderem todo o questionário, os surdos fizeram algumas críticas nessa abordagem. Segundo eles, em alguns momentos durante o filme como, por exemplo, em uma cena de ação, torna-se difícil tirar a atenção da cena e observar alguma tradução no dispositivo.

Tabela 7. Nível de satisfação do surdo em utilizar um dispositivo móvel com a solução proposta.

	Pergunta	Média	Desvio Padrão
Solução Proposta	Foi fácil utilizar o dispositivo móvel?	3,80	1,55
	Foi fácil assistir o filme e acompanhar a tradução no dispositivo móvel ao mesmo tempo?	2,90	1,97

Por fim, alguns resultados puderam ser observados em cada umas das etapas apresentadas nesse experimento. Sendo que, com o objetivo de sintetizá-los, a seção seguinte fará uma análise mais generalizada a respeito desses resultados. Essa análise consiste em fazer um comparativo de desempenho entre as abordagens investigadas.

5.1.3 Análise dos resultados obtidos nas duas etapas do experimento

Os resultados obtidos nesse experimento foram coletados em duas etapas. Em cada etapa, foram investigadas duas abordagens diferentes, onde o objetivo de cada abordagem foi medir o nível de compreensão dos conteúdos exibidos através do filme utilizado nos testes. A Figura 17 apresenta um gráfico onde é possível visualizar os resultados obtidos em cada abordagem investigada.

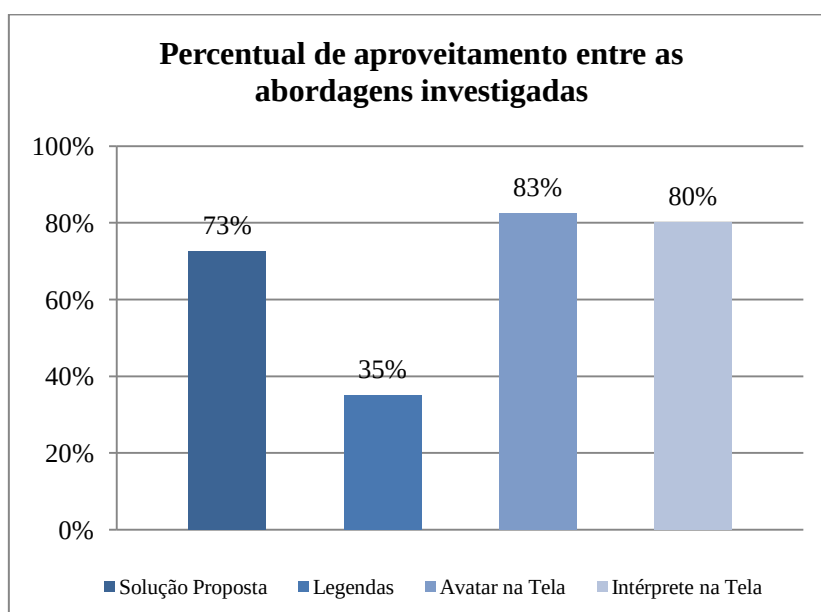


Figura 17. Gráfico do percentual de compreensão do conteúdo exibido no experimento.

O gráfico apresentado na figura acima demonstra que existem indícios de que com o uso de legendas, os usuários surdos tiveram uma taxa média de acertos inferior em relação aos demais tratamentos. Em contrapartida, com o avatar posicionado na tela, os surdos obtiveram 83% de acerto nas questões interpeladas. Com o intérprete posicionado na tela, o aproveitamento foi de 80%, enquanto com a solução proposta, o percentual de acerto foi de 73%. Esses resultados indicam que, em relação ao recurso disponível atualmente, isto é, baseado em legendas textuais, a solução proposta nesse trabalho revelou-se mais eficaz. Em contrapartida, os usuários que acompanharam os conteúdos baseados na solução proposta, apresentaram um desempenho médio próximo dos usuários que acompanharam os conteúdos sobrepostos na tela do filme (Avatar na

Tela e Intérprete na Tela). Isso é um indício que a apresentação dos conteúdos na segunda tela, não prejudica muito a compreensão dos usuários aos conteúdos cinematográficos. Uma análise estatística mais detalhada, no entanto, faz-se necessária para avaliar se essas diferenças são significativas.

Por fim, para melhor visualizar a distribuição dos resultados obtido em cada abordagem, um gráfico de caixa (*box plot*)⁵ também foi construído com esses resultados. Esse gráfico é apresentado na Figura 18.

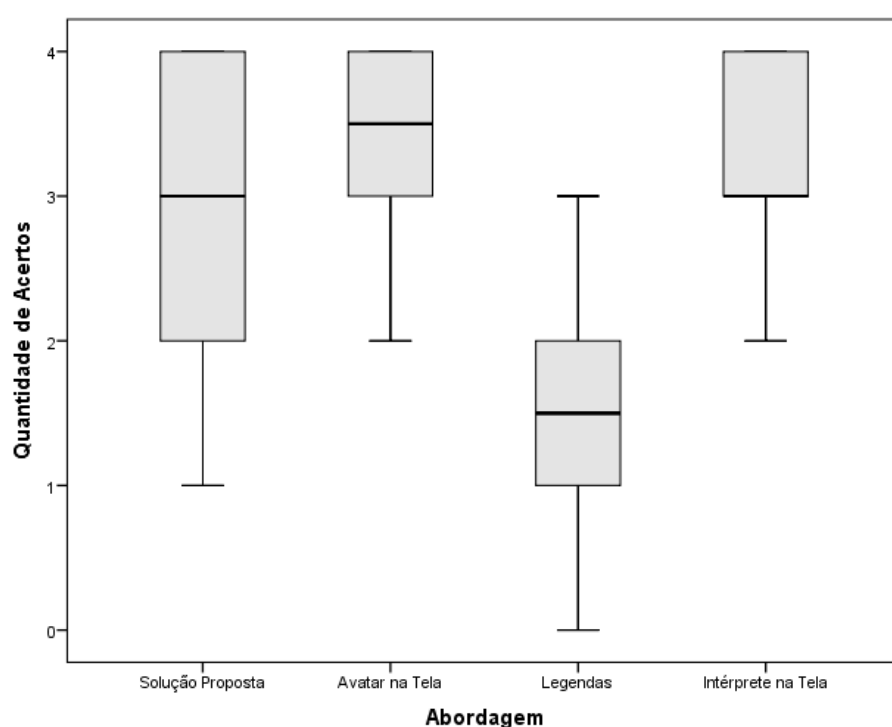


Figura 18. Análise do nível de acerto entre as abordagens através do gráfico de caixa.

Observando o gráfico exposto na figura acima, é possível observar que, quando utilizado o recurso das legendas, o nível de compreensão do conteúdo foi relativamente inferior em relação aos resultados dos demais tratamentos. De modo mais específico, nos testes com legendas, o valor da mediana, primeiro e terceiro quartil foi 1,5, 1 e 2, respectivamente. Isso significa que os surdos que participaram desse teste tiveram uma

⁵ Os gráficos de *box plot* são uma forma de representar resumidamente a dispersão de um conjunto de amostras.

taxa de acerto entre 25% e 50%. Já com os conteúdos baseados na solução proposta, o valor da mediana, primeiro e terceiro quartil foram, respectivamente, 3, 2 e 4.

Fazendo um comparativo entre os resultados obtidos nessas duas abordagens (isto é, com as legendas e com a solução proposta), é possível observar que a solução proposta obteve um nível de aproveitamento superior aos conteúdos com legendas.

Um detalhe importante nos resultados desse experimento que deve ser observado é que, utilizando uma janela com um avatar 3D na tela, os surdos conseguiram uma taxa de acerto um pouco maior em relação aos que utilizaram a janela com um intérprete de LIBRAS. Essa diferença pode ser percebida através dos valores da mediana demonstradas no gráfico da Figura 16. Ou seja, com o avatar 3D na tela, o valor da mediana foi 3,5 (87,5%), enquanto com o intérprete, a mediana foi 3 (75%).

Normalmente, até pela naturalidade dos sinais, os surdos preferem que a tradução dos conteúdos seja feita por um intérprete de LIBRAS. No entanto, os resultados desse experimento apontaram que a abordagem utilizando um avatar 3D posicionando na tela obteve uma média de aproveitamento superior a que utilizou um intérprete. Uma possível explicação para isso pode ser o regionalismo, que dependendo da região, podem existir sinais diferentes para representar a mesma ideia. Ou ainda, pelo fato de não possuir uma amostra significativa para se concluir se uma abordagem é exatamente melhor do que outra.

Conforme mencionado anteriormente, no entanto, uma análise estatística mais detalhada sobre esses fatores se faz necessária para avaliar os motivos desses resultados e para identificar se as diferenças obtidas são significativas ou não. Essa análise estatística detalhada é uma das propostas de atividades futuras deste trabalho.

6 Considerações Finais

Este trabalho apresenta o desenvolvimento de uma solução tecnológica que viabilize a geração automática e a distribuição de janelas de LIBRAS com um avatar 3D nas salas de cinema. A proposta é que essas janelas sejam geradas automaticamente a partir das trilhas de legendas do filme, e distribuídas dentro da sala de cinema utilizando mecanismos para transmissão de conteúdos multimídias. Dessa forma, os surdos poderão utilizar um dispositivo móvel para receber esses conteúdos acessíveis.

Para avaliar o desempenho da solução proposta, foi implementado um protótipo do sistema apresentado no Capítulo 4. Para isso, inicialmente foi definida a arquitetura geral dos componentes, isto é, as responsabilidades inerentes ao modelo conceitual da solução foram separadas módulos. Além disso, para que fosse possível obter os arquivos das trilhas de legenda do filme, foi desenvolvido um protocolo de comunicação entre a solução proposta e o sistema de distribuição de conteúdos do cinema, nesse caso, o Fogo Player.

Com o objetivo de investigar o nível de compreensão dos conteúdos gerados pela solução, foi realizado um experimento com usuários envolvendo o protótipo desenvolvido. Esse experimento avaliou algumas abordagens diferentes, isto é, utilizando as legendas textuais, através de uma janela de LIBRAS com um avatar 3D posicionado na tela, utilizando as janelas com intérpretes de LIBRAS ao invés de avatar 3D, e a solução proposta neste trabalho. Para coletar os dados desse experimento, foi elaborado um questionário onde surdos teriam que respondê-lo após a avaliação de cada abordagem investigada.

Após uma análise dos resultados obtidos nesse experimento, foi possível observar que a solução proposta obteve um desempenho superior comparada ao método convencional, isto é, utilizando o recurso das legendas. No entanto, observando os resultados sobre o nível de facilidade em utilizar um dispositivo móvel para acompanhar a tradução do conteúdo, percebe-se que grande parte dos surdos que participaram do experimento encontraram dificuldades para acompanhar as duas telas simultaneamente. Um ponto colocado pelos surdos a esse respeito é que em alguns

momentos durante o filme como, por exemplo, em uma cena de ação, é difícil tirar a atenção da cena e observar alguma tradução no dispositivo móvel. Mais detalhes sobre esse experimento são apresentados no Capítulo 5.

Sendo assim, considera-se que os objetivos apresentados no Capítulo 1 deste trabalho foram alcançados e que essa solução, de modo geral, obteve resultados bastante significativos. No entanto, é importante ressaltar a necessidade de se realizar novos experimentos e investigações a respeito dessa problemática.

Por fim, apesar de se encontrar em fase de desenvolvimento, é possível perceber a existência de uma contribuição científica, tecnológica e social da proposta apresentada neste trabalho, uma vez que essa solução pode trazer grandes benefícios para aproximadamente 7,5 milhões de surdos brasileiros.

6.1 Trabalhos Futuros

Como sugestão de trabalhos e atividades a serem desenvolvidas posteriormente para dar continuidade a pesquisa apresentada neste trabalho, são recomendados os seguintes pontos:

- *Realizar novos experimentos com uma quantidade maior de usuários surdos.* Embora o experimento realizado neste trabalho tenha fornecido alguns resultados satisfatórios, é importante ressaltar que tratando-se de um público com deficiência auditiva, isto é, pessoas surdas, é importante que sejam realizados uma boa quantidade de testes com esses usuários, pois, dessa forma, será possível validar uma solução viável.
- *Elaborar análises estatísticas mais criteriosas a partir de resultados obtidos em experimentos.* Os resultados apresentados neste trabalho indicam um desempenho superior da solução proposta em relação os recursos convencionais (baseado em legendas de texto), no entanto, não foram feitas análises para demonstrar se a diferença entre o desempenho das abordagens é estatisticamente representativa.

- *Fazer uma revisão sistemática da literatura a respeito do problema abordado neste trabalho.* A pesquisa dos trabalhos relacionados feita neste estudo não seguiu nenhum método científico elaborado. As referências bibliográficas foram exploradas em alguns sistemas de busca como IEEE Xplore⁶, ACM Digital Library⁷, ScienceDirect⁸, Scopus⁹ e Google¹⁰. Portanto, recomenda-se que seja feita uma pesquisa sistemática, baseada em métodos e critérios bem definidos.
- *Realizar testes computacionais com a solução proposta.* Para validar a solução proposta em um cenário real, faz-se necessária a realização de alguns testes voltados para o desempenho computacional da solução. Dessa forma, sugere-se que sejam feitos experimentos com relação ao atraso médio de geração, transmissão, recepção e apresentação das janelas de LIBRAS nos dispositivos dos clientes.

⁶ <http://ieeexplore.ieee.org>

⁷ <http://dl.acm.org>

⁸ <http://www.sciencedirect.com>

⁹ <http://www.scopus.com>

¹⁰ <http://www.google.com>

Referências Bibliográficas

Decreto de Lei nº 5.296. Brasília, 2 de 12 de 2004.

ABNT, Associação Brasileira de Normas Técnicas. “NBR 15290: Acessibilidade em comunicação na televisão.” *Accessibility in tv captions*. Rio de Janeiro, 2005.

Almeida, Jane. “Cinema Digital: novos formatos de expressão e difusão audiovisual.” *Cinemática Brasileira*, 2008.

Amaral, S. T., e M. C. Souza. “O direito de ouvir com os olhos nas TVs brasileiras de sinal aberto.” *Intertemas* 357-386.

“Apple Inc.” 2013. <http://www.apple.com/> (acesso em 2 de Setembro de 2013).

Aquino Júnior, L. L., R. Gomes, M. G. Neto, A. Duarte, R. Costa, e G. L. Filho. “A Software-Based Solution for Distributing and Displaying 3D UDH Films.” *IEEE Computer Society*, Janeiro de 2013: 60-68.

Araújo, S. G., A. C. P. Pedroza, e A. C. Mesquita. “Uma Metodologia de Projeto de Protocolos de Comunicação Baseada em Técnicas Evolutivas.” *XX Simpósio Brasileiro de Telecomunicações*, 5 de Outubro de 2003.

Araújo, T. M. U. “Uma solução para geração automática de trilhas em Língua Brasileira de Sinais em conteúdos multimídia.” Natal, Rio Grande do Norte, 2012. 203f.

Bieler, Rosangela B. “Desenvolvimento Inclusivo: Uma abordagem universal da Deficiência.” Banco Mundial, 25 de Outubro de 2005.

Blakowski, G., and R. Steinmetz. “A media synchronization survey: reference model, specification and case studies.” *IEEE Journal on Selected Areas in Communications*. 1996. 5–35.

BRASIL. *Lei n. 10.098/2000*. Brasília, DF, 19 de Dezembro de 2000.

—. “Acessibilidade.” 2011. <http://www2.planalto.gov.br/acessibilidade> (acesso em 11 de Maio de 2013).

“Brasilmedia.” *Deficientes Auditivos*. 2012. http://www.brasilmedia.com/tipos-de-deficiencia-auditiva.html#Ubs0Y_nVBrt (acesso em Maio de 2013).

BUTTUSSI, F., L. CHITTARO, e M. COPPO. “Using Web3D technologies for visualization and search of signs in an international sign language dictionary.” *12th International Conference on 3D Web Technology*. 2007. 61-70.

- Buttussi, F., L. Chittaro, E. Carchitti, and Marco. Coppo. "Using Mobile Devices to Support Communication between Emergency Medical Responders and Deaf People." 2008.
- CFRF, Conselhos Federal e Regionais de Fonoaudiologia. "Audiometria Tonal, Logaudiometria e Medidas de Imitância Acústica." *Orientações dos Conselhos de Fonoaudiologia para o Laudo Audiológico*. Brasília, 2009 илл Abril.
- Cleary, J.G., and I. H. Written. "Data compression using adaptive coding and partial string matching." *IEEE Transactions on Communications*. 1984. 396–402.
- DCI, Digital Cinema Initiatives - LLC. "Digital Cinema System Specification." 10 de Outubro de 2012.
- "Dicio." 2013. <http://www.dicio.com.br/inteligibilidade/> (acesso em 12 de Julho de 2013).
- FEBRABAN. "Manual de Acessibilidade para Agências Bancárias." *Coleção FEBRABAN de Inclusão Social*. São Paulo, 2010.
- Ferreira, Felipe L. S., Felipe H. Lemos, Gutenberg P. B. Neto, Tiago M. U. Araújo, and Guido L. de S. Filho. "Providing Support for Sign Languages in Middlewares Compliant with ITU J.202." *Proceedings of the 2011 IEEE International Symposium of Multimedia, ISM*, 2011: 149-156.
- Ferreira, Helder F. P. C., e Paulo C. C. Pereira. "HDTV e Cinema Digital." FEUP, 16 de 12 de 2002.
- Filho, Guido L. de Souza, e Tiago M. Ugulino Araújo. "GTAAaS – Acessibilidade como um Serviço (Accessibility as a Service)." João Pessoa, 18 de Agosto de 2011.
- Gallahue, David L., e John C. Ozmun. *Understanding Motor Development: Infants, Children, Adolescents, Adults with PowerWeb*. McGraw-Hill Higher Education, 2005.
- Gallo, B., R San-Segundo, J. M. Lucas, R. Barra, L. F. D'Haro, e F. Fernández. "Speech into Sign Language Statistical Translation System for Deaf People." *Latin America Transactions, IEEE.*, 2009: 400-406.
- Gonçalves, Renata. "Walter Benjamin e a Importância do Cinema na Modernidade." *Existência e Arte - Revista Eletrônica do Grupo PET* (Universidade Federal de São João Del-Rei), 2008.
- Halawani, S. M. "Arabic Sign Language Translation System On Mobile Devices." *IJCSNS International Journal of Computer Science and Network Security*, Vol.8 No.1, Janeiro de 2008: 251-256.

- IBGE. *Instituto Brasileiro de Geografia e Estatística*. 16 de Novembro de 2000.
- IBGE. “Instituto Brasileiro de Geografia e Estatística.” *Censo Demográfico 2000*. <http://www.ibge.gov.br/home/presidencia/noticias/27062003censo.shtm>. 16 de Novembro de 2011.
- “Instituto Passadori.” *A Importância da Comunicação no Processo de Educação*. 2012. <http://www.passadori.com.br/cursos/palestras/artigos/a-importancia-da-comunicacao-no-processo-de-educacao> (acesso em Maio de 2013).
- Kato, Nobuko, Ichiro Naito, Hiroshi Murakami, Hiroshi Minagawa, and Yasushi Ishihara. “Visual communication with dual video transmissions for remote sign language interpretation services.” Ibaraki, Tsukuba: Research Center on Education Media, 2010.
- Lima, Maria do S. C. “Algumas Considerações sobre o Ensino de Português para Surdos na Escola Inclusiva.” *LETRA MAGNA*, 2006: 20-34.
- Luca, Luiz G. A. “A Hora do Cinema Digital.” *Democratização e Globalização do Audiovisual*. São Paulo: Imprensa Oficial, 2009.
- Mascarello, Fernando. “História do Cinema Mundial.” Campinas, SP: Papirus, 2006.
- Menezes, D. C., e T. C. F. Cavalcante. “Compreensão de textos escritos: um estudo com adolescentes surdos.” *Distúrb Comun*, 29 de Abril de 2008: 29-38.
- Moffat, A. “Implementing the ppm data compression scheme.” *IEEE Transactions on Communication*. 1990. 1917–1921.
- Morrissey, S. “Data-driven machine translation for sign languages.” Dublin: Dublin City University, 2008.
- Nielsen, Jakob. *Usability Engineering*. Mountain View: Academic Press, 1994.
- Othman, A., e M. Jemni. “Statistical sign language machine translation: from english written text to american sign language gloss.” *International Journal of Computer Science Issues* 8 (5), 2011: 65-73.
- Pupo, Daise T., Amanda M. Melo, e Sofia P. Ferrés. *Acessibilidade : discurso e prática no cotidiano das bibliotecas*. Campinas, SP: Biblioteca Central Cesar Lattes, 2006.
- San-segundo, R., et al. “Design, Development and Field Evaluation of a Spanish into Sign Language Translation System.” *Pattern Analysis and Applications* 15 (2), 2011: 203–224.
- Silva, Danilo Assis N. dos S. “Uma Linguagem de Definição Expansível para Língua Brasileira de Sinais.” João Pessoa: UFPB, Fevereiro de 2012.

- Silva, F. I., F. Reis, P. R. Gauto, S. G. de L. Silva, e U. Paterno. “Aprendendo Língua Brasileira de Sinais como Segunda Língua.” Santa Catarina: Núcleo de Estudos e Pesquisas em Educação de Surdos, 2007.
- Sintel. 2010. <http://www.sintel.org> (acesso em 19 de Junho de 2013).
- SONY. “Sony Digital Cinema.” *Make believe*. Los Angeles, CA: EUA, 2012.
- Steves, Michelle Potts, Elime Morse, Carl Gutwin, e Saul Greenberg. “A comparison of usage evaluation and inspection methods for assessing groupware usability.” *Proceedings of the 2001 International ACM SIGGROUP Conference on Supporting Group Work*, 2001: 125-134.
- Tambassia, Claudia de A., Ronaldo Rocha, Fernando O. Runstein, and Henrique M. Costa. “Solução para comunicação e interação com deficientes auditivos em sala de aula.” *II ENINED - Encontro Nacional de Informática e Educação* ISSN:2175-5876, 2012: 115-122.
- UNESCO. “Educação: Um Tesouro a Descobrir.” *Relatório para a UNESCO da Comissão Internacional sobre Educação para o século XXI*. Brasília, Julho de 2010.
- Veale, T., B. Collins, e A. Conway. “The Challenges of Cross-Modal Translation: English to Sign Language Translation in the Zardoz System.” *Machine Translation* 13(1), 1998: 81–106.
- VLC. “VLC for iOS.” *iTunes*. 2013. <https://itunes.apple.com/us/app/vlc-for-ios/id650377962?mt=8> (acesso em 2 de Setembro de 2013).
- Wallace, L., e J. Toon. “Case Study: New Wireless Captioning System Debuts at Dallas Cowboys Stadium.” *Georgia Tech Research Institute*. Janeiro de 2010. <http://www.gtri.gatech.edu/casestudy/wireless-captioning-dallas-cowboys-stadium> (acesso em 11 de Janeiro de 2013).
- West, Leanne L., Ethan W. Adler, Jeff Jo, John M. Stewart, and Jack W. Wood. “Display Design Trade-offs for a Wireless Mobile Captioning System.” Georgia Tech Research Institute, 2009.
- Wohlin, C., P. Runeson, M. Höst, M. C. Ohlsson, B. Regnell, e Wesslén A. “Experimentation in Software Engineering: An Introduction.” Norwell, EUA: Kluwer Academic Publisher, 2000.
- Zhao, L., K. Kipper, W. Schuler, C. Vogler, N. Badler, e M. Palmer. “A Machine Translation System from English to American Sign Language.” *Proceedings of the 4th Conference of the Association for Machine Translation in the Americas*. Cuernavaca, 2000. 54–67.

ANEXO A - Questionário aplicado no experimento com os surdos

INFORMAÇÃO USUÁRIO:

SEXO: () MASCULINO () FEMININO

IDADE: _____

VOCÊ É? () SURDO () INTÉRPRETE () INSTRUTOR

COMO VOCÊ SURDO? ESCOLHER:

- () SURDO NASCENÇA
- () TER ALGUÉM FAMÍLIA TAMBÉM SURDO
- () SUA FAMÍLIA USAR LIBRAS SUA CASA
- () VOCE PARTICIPAR COMUNIDADE SURDO?
- () TRABALHA ALGUMA ESCOLA PARA SURDO?
- () ESTUDAR ESCOLA PARA SURDOS?
- () ESTUDAR ESCOLA JUNTO OUVINTE?

VOCÊ SABER LER? () SIM () NÃO

QUAL SEU GRAU ESCOLARIDADE?

- () NUNCA ESTUDAR
- () ENSINO FUNDAMENTAL INCOMPLETO (AINDA NÃO TERMINAR 5º AO 9º ANOS).
- () ENSINO FUNDAMENTAL COMPLETO (TERMINAR 5º AO 9º ANOS).
- () ENSINO MÉDIO INCOMPLETO (AINDA NÃO TERMINAR 1º AO 3º ANOS).
- () ENSINO MÉDIO COMPLETO (TERMINAR 1º AO 3º ANOS).
- () ENSINO SUPERIOR INCOMPLETO (NÃO TERMINAR UNIVERSIDADE).
- () ENSINO SUPERIOR COMPLETO (CONCLUIR CURSO UNIVERSIDADE).
- () OUTRO: _____

ESCOLHER NÍVEL VOCE TER CONHECIMENTO PORTUGUÊS (6 - avançado, 1 - não sei nada) ?

() 6 () 5 () 4 () 3 () 2 () 1

LIBRAS ESCOLHER NÍVEL VOCE TER CONHECIMENTO LIBRAS (6 - avançado, 1 - não sei nada) ?

() 6 () 5 () 4 () 3 () 2 () 1

AVALIAR VÍDEO (SINTEL):

1) FÁCIL ENTENDER? (6 - CLARO); (1 - CONFUSO)

☐ 6 ☐ 5 ☐ 4 ☐ 3 ☐ 2 ☐ 1

2) ACOMPANHA GRAMÁTICA LIBRAS? (6 - PERFEITO); (1- MUITO POUCO)

☐ 6 ☐ 5 ☐ 4 ☐ 3 ☐ 2 ☐ 1

3) SINALIZAÇÃO PARECER NATURAL/NORMAL? (6 - PARECER PESSOA); (1-PARECER ROBÔ)

☐ 6 ☐ 5 ☐ 4 ☐ 3 ☐ 2 ☐ 1

4)MULHER TER ANIMAL ESTIMAÇÃO, QUAL?

☐ GALINHA

☐ GATO

☐ DRAGÃO

☐ SEI-NÃO

5)O-QUE GUARDIÃO FALAR MULHER?

☐ FALAR ELA PERTO ENCONTRAR PRÓPRIO ANIMAL.

☐ FALAR ELA PROCURAR IR CAMINHO ERRADO E ESTAR LONGE.

☐ FALAR ELA SER FILHA DELE.

☐ SEI-NÃO

6) QUAL SER HISTÓRIA LANÇA PERSONAGEM USAR?

☐ LANÇA TER PASSADO TRISTE. MATAR MUITO INOCENTE.

☐ LANÇA TER GLÓRIA PASSADO. MATAR MUITO BANDIDO.

☐ LANÇA PASSADO PERTENCER GRANDE GUERREIRO HÉRCULES PRÓPRIO.

☐ SEI-NÃO

7)O-QUE SIGNIFICAR SINTELL?

☐ SIGNIFICAR TERRA PERIGOSA.

☐ SIGNIFICAR TERRA FRIO

☐ SIGNIFICAR TERRA DRAGÃO

☐ SEI-NÃO

AVALIAR LIBRAS DISPOSITIVO:

1) FÁCIL USAR TECNOLOGIA CELULAR OU TABLET? (6 – MUITO FÁCIL); (1- MUITO DIFÍCIL)

☐ 6 ☐ 5 ☐ 4 ☐ 3 ☐ 2 ☐ 1

2) FÁCIL ACOMPANHAR AVATAR–INTÉRPRETE AO MESMO TEMPO TELA FILME E CELULAR OU TABLET? (6 – MUITO FÁCIL); (1- MUITO DIFÍCIL)

☐ 6 ☐ 5 ☐ 4 ☐ 3 ☐ 2 ☐ 1

ANEXO B - Artigo Publicado no *Journal of Research and Practice in Information Technology*

Accessibility as a Service: Augmenting Multimedia Content with Sign Language Video Tracks

Tiago Maritan U. de Araújo¹, Felipe Lacet Silva Ferreira¹, Danilo Assis Nobre e Silva¹, Eduardo de Lucena Falcão¹, Leonardo Dantas¹, Leonardo de Araújo Domingues¹, Luis Henrique Lopes¹, Yurika Sato¹, Hozana Raquel Lima¹, Alexandre Nóbrega Duarte¹, Guido Lemos S. Filho¹

¹ Digital Video Application Lab, Federal University of Paraíba,
Cidade Universitária, João Pessoa – Paraíba, Brazil
{maritan, lacet, danilo, eduardof, leodantas, leonardo.araujo, luishenrique, alexandre, guido}@lavid.ufpb.br

Abstract: In this paper we explore the concept of “accessibility as a service” by proposing a cloud computing service to help deaf people to access digital contents. The proposed service automatically generates and embeds a sign language video track into multimedia contents. The service organizes the collaboration of sign language experts to dynamically adjust the system that runs on a cloud computing infrastructure. As a case study, we made an implementation of the service for providing support for the Brazilian Sign Language (LIBRAS) and some preliminary tests with Brazilian deaf users to evaluate the proposed solution.

Keywords: accessibility; sign language; machine translation; multimedia contents; cloud computing; sign animation.

Categories and Subject Descriptors: I.2.7 (Natural Language Processing): *Machine Translation*; H.3.5 (Online Information Services): *Web-based services*; H.5.1 (Multimedia Information Systems): *Animations*; H.5.3 (Group and Organization Interfaces): *Collaborative computing*.

1 Introduction

Deaf people have serious difficulties to access information. Sign languages are their primary means of communication, but information and communication technologies (ICTs) rarely provide support for them. In addition, it is difficult for deaf people to understand and communicate by using texts in spoken languages. For example, some reading comprehension tests performed by Wauters (2005) with deaf children aged 7-20 in the Netherlands showed that only 25% of them read at or above the level of a nine-year-old hearing child. In Brazil, about 97% of the deaf people do not finish the high school (IBGE, 2000).

There are several works in the scientific literature developed to address their communication limitations (Lee et al., 2007) (Lee et al. 2005) (Stamer et al., 1998). These works offer technological solutions to daily activities, which enable deaf people to watch and understand television (Lee et al., 2007), to interact with other people (Lee et al., 2005) (Stamer et al., 1998), among others. Others works are related to the use of machine translation methods to translate spoken languages to sign languages (SLs) (Veale et al., 1998) (Zhao et al., 2000) (Morrissey, 2008) (San-Segundo et al., 2008) (Gallo et al., 2009). Veale et al. (1998), for example, described a multilingual translation system used for translating English texts into Japanese Sign Language (JSL), American Sign Language (ASL) and Irish Sign Language (ISL).

Zhao, et al. (2000) developed an interlanguage-based approach for translating English text into ASL. Input data are analyzed and an intermediate representation (IR) is generated from their syntactic and morphological information. Then, a sign synthesizer uses the IR information to generate the signs. Morrissey (2008) proposed an example-based machine translation (EBMT) system for translating text into ISL. To do this task, an EBMT approach needs a bilingual corpus. However, due to the lack of a formally adopted or recognized writing system

Figura 19. Artigo Publicado no *Journal of Research and Practice in Information Technology* sob o tema: *Accessibility as a Service: Augmenting Multimedia Content with Sign Language Video Tracks*.

ANEXO C - Artigo Publicado na II Escola Paraibana de Informática

Uma Experiência de Avaliação de Interface de Usuário em Dispositivos “Segunda Tela” no aplicativo CineLIBRAS

Cecilia Flávia da Silva, Francielly Cardoso Grigorio, Sérgio Dias Carvalho Silva, Felipe Lacet, Leonardo Domingues, Tatiana Tavares, Tiago Maritan

Universidade Federal da Paraíba (UFPB)

Campus Universitário Castelo Branco – João Pessoa– PB – Brasil

cecilia@lavid.ufpb.br, francielly@lavid.ufpb.br, sergiodias@di.ufpb.br, {lacet, leonardo, tatiana, tiagomaritan}@lavid.ufpb.br

Abstract. The present paper describes the evaluation of interfaces in “second display devices”, which consists in the use of additional devices to increment the interaction with users. Particularly in this case, the second display device evaluated is intrinsically related with accessibility issues. Thus, the evaluation is based in study the performance of the interface applied in a target public - with hearing deficiency. For this purpose, it was utilized the project application CINELIBRAS, developed by Digital Video Applications Laboratory (LAViD).

Resumo. O presente artigo descreve a avaliação de interfaces em dispositivos “Segunda Tela”, termo que consiste no uso de dispositivos adicionais para incrementar a interação com o usuário. Particularmente neste caso, o dispositivo segunda tela avaliado está intrinsecamente relacionado com acessibilidade. Assim, a avaliação baseia-se em estudar o desempenho da interface aplicada em um público-alvo - com deficiência auditiva. Para este fim, foi utilizada a aplicação do projeto CINELIBRAS, desenvolvida pelo Laboratório de Aplicações de Vídeo Digital (LAViD).

1. Introdução

Interação Humano-Computador (IHC) é a área da computação que se preocupa com o design, avaliação e implementação de sistemas computacionais interativos para uso humano [Hewett, 1992]. Esta interação com o humano é feita através da interface, responsável por apresentar um conjunto de ferramentas que auxiliam o mesmo na realização de suas tarefas.

Entretanto, não é sempre que o usuário consegue interagir com a interface sem dificuldades, tomando necessário a execução de um processo de avaliação que pode permear o processo de desenvolvimento do software. Esse fato ainda é mais crítico quando falamos de interfaces acessíveis, ou acessibilidade. Para atender requisitos relacionados à acessibilidade, a interface precisa ser elaborada com atenção reforçada, visto que a interface terá a responsabilidade de apresentar seu conteúdo para usuários com necessidades especiais.

Figura 20. Artigo publicado na II Escola Paraibana de Informática com o tema: “Uma Experiência de Avaliação de Interface de Usuário em Dispositivos “Segunda Tela” no aplicativo CineLIBRAS”.

ANEXO D - Prêmio de Melhor Artigo de Iniciação Científica na II Escola Paraibana de Informática



Figura 21. Prêmio de melhor artigo de iniciação científica pelo artigo intitulado como: "Uma Experiência de Avaliação de Interface de Usuário em Dispositivos "Segunda Tela" no aplicativo CineLIBRAS"