

# COMBINANDO ABORDAGENS NA VISUALIZAÇÃO DE DADOS EM UMA FERRAMENTA DE CARTOGRAFIA LINGUÍSTICA

## COMBINING APPROACHES IN DATA VISUALIZATION IN A LINGUISTIC CARTOGRAPHY TOOL

Rodrigo Duarte Seabra, Nathan Oliveira, Valter Pereira Romano  
rodrigo@unifei.edu.br, oliveiranathan@gmail.com, valter.romano@hotmail.com  
Universidade Federal de Itajubá/Brazil

**Abstract:** Computer Graphics and Image Digital Processing algorithms have been increasingly exploited in various areas of the human knowledge, such as Medicine, Engineering, Games, among others. Considering the evolution of computer technologies currently available with respect to the computational power, these algorithms mostly aims to amplify the experience of user interaction, to privilege aspects of information visualization and, in some cases, to increase the sense of presence. However, techniques designed often involve calculations in real time and realistic renderings that require the development of high-performance algorithms. This paper presents aspects of the implementation of algorithms developed to support to the data visualization into a tool for generating and visualizing linguistic maps, especially as regards the calculation of completing and area intersection detection in isogloss maps and generating gradients to represent the lexical variant occurrence productivity in these maps. Test conducted in generating of gradients showed that use of a parallel implementation achieved impressive results in the algorithm performance.

**Keywords:** Algorithms; Data visualization; Linguistic cartography.

**Resumo:** Algoritmos de Computação Gráfica e de Processamento Digital de Imagens têm sido cada vez mais explorados nas mais variadas áreas do conhecimento humano, a exemplo de aplicações na Medicina, Engenharia, Jogos, dentre outras. Considerando a evolução das tecnologias computacionais atualmente disponíveis no que tange ao poder computacional de processamento, esses algoritmos, em sua maioria, visam a amplificar a experiência de interação do usuário, privilegiar aspectos de visualização da informação e, em alguns casos, aumentar a sensação de presença. Todavia, as técnicas concebidas frequentemente envolvem cálculos em tempo real e renderizações realísticas que requerem o desenvolvimento de algoritmos com alto desempenho. Esse trabalho apresenta aspectos de implementação de algoritmos desenvolvidos no apoio à visualização de dados em uma ferramenta para geração e visualização de cartas linguísticas, sobretudo no que se refere ao cálculo de preenchimento e detecção de interseções de áreas em cartas isoglossas e na geração de gradientes de tonalidades para representar a produtividade das ocorrências das variantes lexicais nessas cartas. Testes realizados na geração dos gradientes mostraram que o uso de uma implementação paralela alcançou resultados expressivos no desempenho do algoritmo.

**Palavras-chave:** Algoritmos; Visualização de dados; Cartografia linguística.

### I. INTRODUÇÃO

O interesse pela variação linguística e a sua representação em mapas geográficos remonta ao fim do século XIX e tem como figura precursora o alemão George Wenker (1851-1911), cujo trabalho inaugural acerca das diferenças fonéticas na língua germânica chamou a atenção dos estudiosos a um novo modo de observação das línguas, representando-as em mapas geográficos.

Desse modo, pode-se registrar a área de abrangência de determinada língua e/ou dialeto, podendo-se fazer a comparação entre formas linguísticas diferentes, bem como explicar as interferências de formas distintas. Porém, foi com o trabalho do francês Jules Gillieron e seu assistente Edmond Edmont [1] que o método da geografia linguística, modernamente chamada de geolinguística, veio a se consolidar no bem sucedido *Atlas linguistique de la France* (1915).

Tendo, portanto, como cenário inicial, a Europa da primeira metade do século XX, a geolinguística alcançou outros continentes, refletindo-se na América com Hans Kurath (1939) na elaboração do *Linguistic Atlas of New England* (LANE) [2] e Tomás Navarro (1946) [3], acerca do espanhol porto-riquenho, além de outros estudiosos que os sucederam na Colômbia, no México, no Chile e no Brasil. Neste último, de acordo com Romano [4], a geolinguística tem alcançado inúmeros avanços, desde a publicação do primeiro atlas linguístico de um estado brasileiro, o *Atlas Prévio dos Falares Baianos* [5].

Todo avanço alcançado pela geolinguística brasileira, contudo, tem exigido investimentos e aperfeiçoamentos nas diversas etapas da pesquisa, no sentido de agilizar e otimizar o trabalho dos estudiosos, seja em ferramentas audiovisuais para a coleta de dados, seja na criação de banco de dados para organização do material linguístico coletado. Nesse sentido, a ferramenta [SGVCLin] - Software para Geração e Visualização de Cartas Linguísticas, desenvolvida no âmbito de uma pesquisa de doutorado na área de Dialectologia, contribui para as pesquisas desta natureza, não somente por armazenar o material linguístico em um banco de dados de forma fácil e segura, mas, também, por possibilitar o acesso ao material em forma de relatórios diversos e cartas linguísticas de produtividade e de isoglossas [6].

Neste contexto, e focando aplicações envolvendo o uso de técnicas de processamento digital de imagens e computação gráfica, este trabalho apresenta aspectos de implementação de algoritmos desenvolvidos no apoio à visualização de dados representados em cartas linguísticas por meio da ferramenta supramencionada. A contribuição principal consiste em apresentar detalhes relativos ao cálculo de preenchimento e detecção de interseções de áreas em cartas de isoglossas inspirados em [7], e na geração de gradientes de tonalidades para representar a produtividade das ocorrências das variantes lexicais nessas cartas com base em [8], executando o algoritmo em uma implementação paralelizada.

Na próxima seção, discute-se o referencial teórico sobre o processo de armazenamento de dados e cartografia linguística, bem como aplicações de computação gráfica e processamento de imagens utilizadas nesta pesquisa, mais especificamente no que se refere à detecção de regiões, bordas e interpolação de dados em duas dimensões. Na seção 3, são apresentadas as funcionalidades que compõem a ferramenta de cartografia linguística responsável pelo armazenamento, geração e visualização dos dados. A seção 4 descreve os detalhes de desenvolvimento e aplicação dos algoritmos utilizados para a visualização dos dados linguísticos gerenciados pela ferramenta, bem como os princípios envolvidos no experimento conduzido para a medição do desempenho de uma implementação paralela de um dos algoritmos apresentados. Finalmente, a seção 5 discorre sobre as conclusões e trabalhos futuros.

## II. REFERENCIAL TEÓRICO

### A. Cartografia Linguística

Para o dicionário Houaiss [9], o verbete cartografia apresenta duas acepções “(1) conjunto de estudos e operações científicas, técnicas e artísticas que orienta os trabalhos de elaboração de cartas geográficas; (2) descrição ou tratado sobre mapas”. Desse modo, o termo pode ser aplicado a qualquer operação de base científica que visa a representar os resultados obtidos sobre mapas geográficos e, por extensão, pode ser aplicado, nesta concepção ampla do termo, aos estudos linguísticos.

Há tempos os estudiosos da linguagem têm representado a diversidade das línguas em “mapas especiais” [10]. Assim, de acordo com objetivos previamente estabelecidos, depois de levantado o *corpus* oral de determinada comunidade, devidamente transcrito e organizado, o trabalho de natureza geolinguística passa para o processo de cartografia [11].

No cenário nacional, observa-se que, em geral, a etapa de cartografia linguística, passo importante para a construção do atlas linguístico, é realizada, em sua maioria, por profissionais ou da área da Geografia, que possuem conhecimentos específicos de cartografia e de SIG (Sistemas de Informação Geográfica) ou por *designers gráficos*, pessoas habilitadas a trabalhar com

*softwares* de edição de imagens. Raramente o próprio linguista produz suas cartas devido a dois motivos fundamentais: (i) pelo grande conjunto de dados a ser analisado e estudado, que exige amplo investimento de tempo e (ii) pela falta de conhecimentos computacionais de *softwares* de edição e geração de imagens.

Nesse sentido, o linguista vê a necessidade de “terceirizar” a cartografia de seu atlas, ficando responsável em fazer a exegese do *corpus* e transferir as informações necessárias para o profissional que representará espacialmente o material linguístico selecionado. Ou seja, a carta linguística é elaborada sem a participação direta do linguista, que somente tem acesso ao produto final – a carta pronta – ficando responsável pela revisão e, quando necessário, solicita os ajustes ao profissional contratado.

O cartógrafo ou *designer* é quem realiza as tarefas, constituindo-se como mediador entre o material linguístico representado na carta e o linguista. Este processo de intermediação torna a produção do atlas linguístico uma atividade morosa, que acaba arrastando o projeto do atlas por anos a fio até que se consiga obter um número considerável de cartas linguísticas para a publicação.

Esta tem sido a solução imediata encontrada pelos geolinguistas brasileiros para produção dos atlas. No entanto, observa-se que a elaboração de uma carta linguística é uma tarefa complexa, pois, em consonância às ideias de [12], “fazer uma carta não é mero trabalho de transposição das respostas fornecidas para um cartograma”. A cartografia linguística exige requisitos que vão além das habilidades técnicas. São necessários conhecimentos intelectuais próprios daqueles que trabalham neste campo de estudo, uma vez que, além de saber como representar, é necessário saber o que representar, quais variantes são válidas, qual extensão de legenda, a natureza da carta (léxica, fonética, morfosintática, isoglósica), o tipo de representação, dentre outros fatores, que o “não-linguista”, em geral, não compreende e, às vezes, ignora por priorizar aspectos estéticos ou mesmo conceituais de outras áreas do saber.

Como já destacado anteriormente, o processo de seleção e exegese dos dados cabe ao linguista, porém, se ele mesmo conseguir gerar suas próprias cartas linguísticas de forma automatizada, sem o envolvimento e dependência de terceiros, o trabalho torna-se mais produtivo. Com isso, é possível cartografar quantos itens achar necessários, bem como excluir de seu atlas as cartas que foram produzidas que, no entanto, não são significativas para o estudo, sem perda de tempo ou investimento financeiro.

Observa-se que a cartografia linguística dos primeiros cinco atlas linguísticos brasileiros publicados, o *Atlas Prévio dos Falares Baianos* – APFB [5], o *Esboço de um Atlas Linguístico de Minas Gerais* – EALMG [13], o *Atlas Linguístico da Paraíba* – ALPB [14], o *Atlas Linguístico de Sergipe* – ALS

[15], e o *Atlas Linguístico do Paraná* – ALPR [16] foi elaborada de modo “manual”, sobretudo o primeiro.

Todavia, atualmente, observa-se que os avanços tecnológicos das últimas décadas tem propiciado a criação de projetos de atlas que consideram o potencial da informática, sobretudo na estruturação de bancos de dados geolinguísticos e na cartografia computadorizada.

No Brasil, os geolinguistas já têm dado os primeiros passos em considerar os avanços tecnológicos na elaboração de atlas linguísticos. Um exemplo disso são os denominados atlas de terceira geração, conforme a classificação de [17].

Embora ainda não estejam difundidos os atlas dessa categoria, é inegável a tendência já inaugurada pelo ALiSPA [18], que, ao considerar recursos computacionais no processo de cartografia e apresentação dos dados, demonstra um grande avanço da Geolinguística brasileira. Contudo, a grande maioria dos trabalhos geolinguísticos não tem incrementado estas novas tecnologias pelos motivos já apontados.

Quanto à organização de bancos de dados geolinguísticos, já se encontram disponíveis algumas opções, porém desenvolvidas sob a perspectiva mercadológica, não necessariamente científica. Ou seja, profissionais da área da Computação desenvolvem ferramentas, *desktop* ou *online*, que permitem ao usuário armazenar os dados geolinguísticos para consultas diversas por meio de relatórios. Exemplos de *softwares* que geram relatórios diversos são o Sistema *Agium Search*, desenvolvido e utilizado em trabalhos como o de [19] e [20], e o *SGP (Sistema Geral de Pesquisa)* utilizado por [21]. Essas ferramentas mostram-se proveitosas, uma vez que otimizam o tempo do pesquisador. Porém, apresentam um agravante no que se refere à manutenção e ao custo, que, embora não seja elevado, requer a intervenção de um terceiro na atualização e manutenção do servidor.

A discussão sobre a importância de gerenciamento de banco de dados geolinguísticos não é tão recente. Já no final da década de 1990, [22] discutia a importância da informatização de dados geolinguísticos apontando para possíveis soluções que otimizassem o tempo do pesquisador e agilizassem a publicação do atlas linguístico.

Foi a partir do programa computacional desenvolvido por Vieira e sua equipe – o *Sistema de Processamento de Dados Geolinguísticos (SPDGL)* – que o primeiro atlas linguístico brasileiro totalmente informatizado foi publicado – o *Atlas Linguístico-Etnográfico da Região Sul* – ALERS [23]. Este mesmo programa foi utilizado também por [24], com adaptações, para a organização e cartografia dos dados do *Atlas Linguístico do Mato Grosso do Sul* – ALMS. Além desses dois atlas, não se tem notícia de outros projetos que utilizaram o SPDGL para armazenamento dos dados e geração das cartas. Outro atlas linguístico estadual que utilizou um sistema informatizado em sua

cartografia foi o *Atlas Linguístico do Amazonas* – ALAM [25].

Entretanto, observa-se que as soluções encontradas para o armazenamento e cartografia linguística como a do ALERS, a do ALMS e do ALAM, ou mesmo as propostas de [22] resolvem problemas específicos de cada projeto e não são divulgadas à comunidade científica, haja vista que o banco de dados é modelado de acordo com as especificidades de cada trabalho. Em suma, o que se tem notado é que cada projeto tem empregado metodologias diferenciadas no armazenamento e cartografia dos dados linguísticos, seja com programas computacionais vinculados a banco de dados, seja utilizando programas de edição de imagens.

### B. Abordagens de Computação Gráfica e Processamento Digital de Imagens

As imagens digitais encontram-se difundidas em diversas aplicações e isso se deve ao fato de que grande parte das informações que o ser humano obtém do mundo em que se encontra inserido ocorre por meio de imagens disponíveis em uma diversidade de objetos, situações e meios de comunicação. Segundo Conci, Azevedo e Leta [26], a área de processamento de imagens abrange tópicos como diminuição de ruídos, realce e restauração de imagens e recuperação de imagens armazenadas em banco de dados, com base no conteúdo visual. Caso haja a necessidade de interpretação do conteúdo da imagem, ela é de responsabilidade do usuário. Todavia, a interpretação de informações presentes em imagens também pode ser realizada pela aplicação de algoritmos computacionais.

Pedri e Schwartz [27] destacam alguns exemplos de domínios do conhecimento que envolvem o uso de técnicas de processamento de imagens na resolução de problemas, a saber: medicina, biologia, automação industrial, sensoriamento remoto, astronomia, microscopia, artes, área militar, arqueologia, segurança e vigilância. Com base nos exemplos citados, nota-se que os campos de aplicação que se beneficiam do processamento de imagens perpassam de áreas que vão da saúde ao monitoramento.

Em adição aos exemplos supramencionados, acrescenta-se a geolinguística, que visa a estudar as línguas em seu contexto geográfico. Essa área interdisciplinar, compartilhada pela linguística e a geografia, tem como duas de suas principais tarefas identificar e descrever áreas caracterizadas por seus domínios linguísticos e dialetais, determinando regiões de influência – em função de diferenças linguísticas no território analisado – representadas pela distribuição de variantes linguísticas em um mapa composto por uma rede de pontos (localidades).

No que tange aos aspectos computacionais, a rede de pontos e as áreas de influência de cada variante linguística em determinadas regiões podem ser analisadas a partir de técnicas de representação que realizem agrupamentos de pontos em uma superfície

(neste caso, o mapa) a partir de um grupo de pontos iniciais. A tecelagem de Voronoi [28, 29] se apresenta como uma solução, pois, nesta técnica, pontos são agrupados de forma que cada grupo contém apenas uma semente e, cada ponto desses grupos, esteja mais próximo da semente do grupo ao qual pertence se comparado às outras sementes. Assim, pontos apresentando propriedades similares são agrupados para formar uma região. Sugerem-se os estudos de [30, 31, 32] para revisões detalhadas sobre a caracterização de regiões em imagens em função de variadas propriedades.

Prosseguindo com essa abordagem, as regiões apresentam limites representados por suas bordas. As bordas são regiões da imagem caracterizadas por uma mudança brusca do nível de iluminação ou tom de duas regiões. Um contorno é uma linha fechada formada pela união de duas bordas de um objeto [26]. Na prática, a detecção de bordas é uma forma de se obter as características de uma imagem por meio de um processamento matemático executado sobre as informações numéricas contidas em cada ponto da imagem. Neste caso, a transição entre duas regiões pode ser determinada com base apenas na descontinuidade dos níveis de cinza. Um dos métodos de detecção de bordas é a convolução de uma matriz especial – o *kernel* – sobre a imagem. Nesta pesquisa, foi utilizado o *kernel* do Laplaciano da Gaussiana [7, 33, 34, 35, 36], de forma a se detectar as bordas que definem as regiões dos mapas com o mínimo de ruído possível.

Finalmente, a interpolação de dados em 2D consiste em um modo de se obter uma superfície contínua, de forma a se calcular suas curvas de nível. No estudo é utilizado o método de Shepard [8] no cálculo de gradientes nas cartas de isoglossas, pois fornece um algoritmo rápido e que pode ser facilmente paralelizado. O motivo de aplicar a paralelização se justifica pelo fato de apesar de possuir alto desempenho, para imagens com alta resolução, o tempo de execução do algoritmo ainda é elevado.

### III. ASPECTOS DE IMPLEMENTAÇÃO DO [SGVCLIN]

Para a definição dos requisitos da ferramenta foram utilizadas variadas cartas linguísticas com dados representativos de amostras reais, com distintas representações, além dos conhecimentos discutidos em reuniões subsidiadas pelo Projeto ALiB – regional Paraná [37].

Durante a etapa de levantamento dos requisitos, foram analisados dois tipos principais de cartas linguísticas a serem geradas: (i) cartas nas quais os dados são representados por gráficos em formato pizza (*pie chart*) posicionados em cada ponto geográfico, e (ii) cartas nas quais são mostradas as áreas de influência de cada variante linguística. A primeira categoria foi de simples implementação, enquanto a segunda necessitou de estudos detalhados que indicaram a tecelagem de Voronoi [28, 29] como melhor alternativa para a aproximação dos cálculos das áreas de influência. Em termos práticos, esta

técnica permite a subdivisão das regiões em um conjunto de áreas de abrangência estabelecendo relações de proximidade.

No que diz respeito às tecnologias utilizadas, os dados são armazenados em um banco de dados MySQL e a linguagem adotada para o desenvolvimento da ferramenta foi o Java. De modo geral, o sistema consiste em uma ferramenta computacional projetada para fornecer uma interface intuitiva e simples, que permita ao usuário o armazenamento de dados geolinguísticos e posterior geração de cartas linguísticas por meio de consultas ao banco de dados. A ferramenta ainda proporciona a possibilidade de criação de grupos de questões por usuário e a geração de relatórios pautados por diferentes variáveis, bem como a criação de cartas bidimensionais: diatópica/diassexual, diatópica/diageracional [38], além de cartas de isoglossas.

Neste momento, vale destacar que as cartas bidimensionais representam a variação linguística considerando duas dimensões: (i) o espaço geográfico (dimensão diatópica); e (ii) o gênero (dimensão diassexual) ou a faixa etária (dimensão diageracional). Já as cartas de isoglossas representam as áreas de abrangência de uma ou mais variantes em uma região geográfica.

Na ferramenta [SGVCLin], a variação das cartas de isoglossas com uso de gradiente provê um modo de visualização que permite ao usuário identificar a quantidade de ocorrências de uma determinada variante, em cada ponto geográfico, por meio de uma exibição elegante.

Por se tratar de uma ferramenta para uso no âmbito da pesquisa e como recurso complementar no ensino da cartografia de dados linguísticos, o sistema utiliza uma interface convencional em conformidade com o padrão WIMP. A interação do usuário com o sistema ocorre via interface, sem que qualquer manipulação direta ao banco de dados seja necessária.

A motivação principal que direcionou o desenvolvimento do *software* foi a criação de um ambiente no qual o linguista possa ter independência no manuseio da ferramenta. Entretanto, pelo fato de a Geolinguística estar inserida na interface da Linguística e da Geografia, cabe ao estudioso, antes de iniciar o uso do *software*, contatar um geógrafo para que este prepare sua base cartográfica ou carta base, que esteja devidamente georreferenciada, contendo os elementos indispensáveis, tais como a escala gráfica, os limites políticos do território e a localização dos pontos linguísticos. Uma vez elaborado o mapa em formato de imagem (extensões ‘png’, ‘jpg’ ou ‘bmp’), o linguista poderá carregá-lo na ferramenta e, assim, representar inúmeros fatos linguísticos (fonéticos, lexicais, morfossintáticos).

As características da interface da ferramenta são sumarizadas a seguir:



- Manipulação de arquivos: a base de dados armazena novas informações na medida em que elas são geradas, requerendo apenas uma conexão ativa. Uma nova base de dados pode ser recriada, excluindo informações salvas. Ademais, a base de dados pode ser exportada para um arquivo, permitindo sua cópia. Arquivos exportados podem ser carregados na ferramenta e seus dados são copiados para a base de dados;
- Impressão de cartas linguísticas e relatórios: mapas linguísticos gerados para arquivos de imagens. Os relatórios são salvos em formato PDF. Usuários podem enviar cartas e relatórios salvos para impressão;
- Questionários: categorias e subcategorias podem ser criadas pelos usuários, bem como eles podem manipular simultaneamente diferentes tipos de questões agrupadas em áreas semânticas em um ou mais questionários;
- Questões e variantes: permite a criação de grupos de questões personalizadas e identificadas individualmente por número e formulação (título). Em seguida, o usuário insere a lista de variantes desejadas para a questão;
- Manipulação e edição de imagens: mapas armazenados em formato de imagem podem ser carregados e a ferramenta ainda pode trabalhar com mais de um mapa base. Pontos representando localidades podem ser registrados sobre os mapas por meio de cliques do mouse. A preparação da base cartográfica para a geração de cartas de isoglossas ocorre via um editor de mapas, e o usuário indica as regiões da imagem que serão consideradas no cálculo das áreas;
- Edição da rede de pontos: possibilita a edição de informações da rede de pontos registrada (nome da localidade e unidade federativa);
- Fichas dos informantes: estabelece o modelo de ficha do informante para ser preenchida. O modelo é flexível e os usuários podem criar campos além dos já existentes. Informações sobre gênero, idade, estado civil, nível educacional, profissão, religião e classe social são a base para a geração de variados relatórios;
- Registro de fichas: informações pessoais dos informantes podem ser registradas a partir do modelo de ficha previamente definido;
- Registro de respostas: é permitido o registro das respostas dos informantes em cada localidade. Após selecionar a questão desejada, o usuário inclui as respostas rapidamente pelo clique do mouse em uma tabela. Para acelerar o preenchimento, as variantes já registradas na etapa de criação das questões e variantes são exibidas em uma lista *drop-down*. Quaisquer ocorrências de informantes que não souberam responder ou no caso de detecção de problemas técnicos também podem ser registradas;
- Exportação de dados: geração de arquivos em formato de planilha com todas as informações fornecidas pelos usuários. Os arquivos gerados são usados como material para tratamentos estatísticos detalhados;
- Relatórios: geração de relatórios distribuídos em treze diferentes formatos de acordo com critérios estabelecidos pelo usuário. As informações são apresentadas por questão e registradas em função das respostas em ordem decrescente de produtividade;
- Cartas: as cartas geradas são diatópica, diatópica/diassexual, diatópica/diageracional e isoglossa. Pares de equivalência indicando a mesma resposta podem ser estabelecidos. A quantidade de variantes lexicais exibidas nos mapas é definida pelo usuário;
- Visualização de mapas: a distribuição das respostas dos informantes nas cartas é exibida em gráficos (em formato pizza) sobre o mapa na interface, além de uma legenda e de um histograma com a produtividade geral;
- Manipulação de gráficos e mapas: os gráficos podem ser selecionados e suas posições transladadas livremente pelo uso das teclas direcionais do teclado do usuário. Isso permite melhor ajuste durante o processo de criação da carta. A seleção da carta segue o mesmo princípio e oferece a possibilidade de explorar regiões da carta que possivelmente não estejam visíveis em determinado momento da interação;
- *Zoom*: o usuário possui três opções de aplicar o efeito de *zoom* para interagir com as cartas, usando o mouse, o teclado ou um menu gráfico representado por setas e uma lupa no canto superior esquerdo da janela. Os gráficos podem ter seu diâmetro alterado usando as teclas '+' e '-'.

#### IV. DESENVOLVIMENTO E APLICAÇÃO DOS ALGORITMOS

Nesta seção, são apresentados detalhes de desenvolvimento e aplicação dos algoritmos utilizados para a visualização dos dados linguísticos representados nas variadas modalidades de cartas geradas pela ferramenta [SGVCLin].

##### A. Cálculo da Área de Preenchimento das Cartas Isoglossas

As cartas isoglossas são traçadas a partir de dois conjuntos de dados: (i) os pontos, inseridos pelo usuário durante a etapa de criação do mapa em que se encontram as localidades; (ii) a área tratada do mapa que deve ser preenchida com as cores das variantes, neste caso, as áreas que não fazem parte das localidades analisadas não são preenchidas. Com o objetivo de permitir a seleção da área a ser preenchida, foi desenvolvida uma interface (Fig. 1) que permite ao usuário indicar as áreas a serem consideradas no cálculo de maneira semiautomatizada.

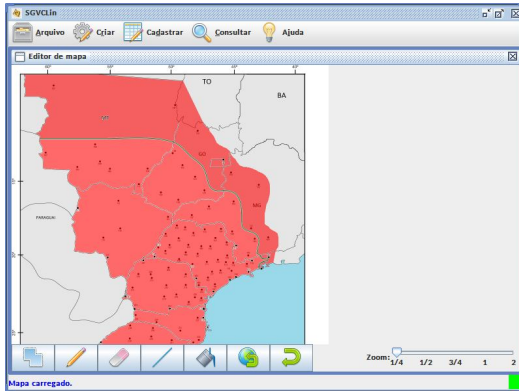


Fig. 1. Interface do editor de mapa. O usuário determina as áreas (no exemplo, indicadas em vermelho) a serem consideradas no cálculo das isoglossas.

A partir da interação do usuário por meio de funcionalidades disponíveis no editor, a interface fornece a opção de detecção automática das bordas sem que o usuário necessite realizar o traçado manual das bordas da imagem desejada. Essa funcionalidade foi projetada devido à natureza dos mapas, que raramente são compostos apenas por linhas retas. Para a detecção de bordas foi implementado um detector por meio da convolução sobre a imagem transformada em tons de cinza, aplicando o *kernel* do Laplaciano da Gaussiana (Fig. 2), e selecionados os pontos que obtiverem valor positivo na aplicação.

-1	-1	-1
-1	8	-1
-1	-1	-1

Fig. 2. *Kernel* do operador Laplaciano da Gaussiana. Por meio da convolução da matriz sobre a imagem, suas bordas são detectadas.

Após o usuário selecionar a área a ser considerada no cálculo, o [SGVCLin] divide esta área usando as informações dos pontos inseridos pelo usuário anteriormente. Para o cálculo das áreas de influência de cada localidade é realizada a divisão da área selecionada (representada em vermelho na Fig. 1) pelo usuário, utilizando os pontos de localidades inseridos, de forma que cada ponto da área selecionada esteja em uma das áreas de localidade, sendo que a área a qual ele pertence é determinada pelo Algoritmo 1, Fig. 3. Após esse cálculo, os dados são armazenados no banco de dados para que não haja a necessidade de serem recalculados.

##### Algoritmo 1: Cálculo das Áreas de Preenchimento

Seja  $w$  a largura da imagem e  $h$  a altura da imagem,  
Seja  $T$  o conjunto de todos os pontos  $(x,y)$  da área a ser preenchida,

Seja  $P$  o conjunto de todos os pontos  $(x,y)$  representando as localidades,

Para todo  $P_i$  pertencente a  $P$ :

Seja  $R$  um conjunto de retas,  $R \leftarrow \{\}$ ,

Para todo  $P_j$  pertencente a  $P$ ,  $P_j \neq P_i$ :

$R \leftarrow R \cup \{P_i P_j\}$ .

Seja  $V$  um conjunto de pontos,  $V = \{(0,0),(0,h),(w,h),(w,0)\}$ ,

Para todo  $R_k$  pertencente a  $R$ :

Seja  $V_n$  um conjunto de pontos,  $V_n \leftarrow \{\}$ ,

Seja  $r$  uma reta perpendicular a  $R_k$  e contendo o ponto central de  $R_k$ ,

Para todo  $V_i$  pertencente a  $V$ :

Seja  $Q$  um conjunto de pontos,  $Q$

$\leftarrow P_i V_i \cap r$

Se  $Q = \emptyset$  ou  $|Q| = 1$ ,

Então:

$V_n \leftarrow V_n \cup \{V_i\}$

Senão:

$V_n \leftarrow V_n \cup Q$

Seja  $A$  um conjunto de pontos contendo todos os pontos internos ao polígono definido pelo conjunto de vértices  $V$ ,

Determinamos  $R$  como sendo a área pertencente ao ponto  $P_i$ ,  $R \leftarrow A \cap T$

Fig. 3. Algoritmo Cálculo das Áreas de Preenchimento

##### B. Cálculo das Interseções nas Isoglossas

Durante a implementação das isoglossas, foi necessário limitar o número máximo de variantes (quatro) que o usuário pode selecionar. Isso se deve ao cálculo das interseções entre as variantes, que, no caso de quatro selecionadas, pode-se alcançar até quinze grupos diferentes, deixando extensas as legendas das cartas. Além disso, caso não houvesse essa restrição, aumentar-se-ia a possibilidade de ser necessário o uso de cores semelhantes nas cartas, o que diminuiria sua legibilidade.

Para o cálculo dos grupos de variantes que se interceptam, mas não são vazios, foi usado um algoritmo simples (Algoritmo 2, Fig. 4) que determina,

para cada ponto representando uma localidade, quais variantes selecionadas ocorrem no ponto, e seleciona uma cor aleatória para cada grupo diferente.

**Algoritmo 2: Cálculo das Interseções nas Isoglossas**

Seja  $q$  a quantidade de variantes selecionadas,  
 Seja  $G$  um conjunto de conjuntos de localidades,  $|G| = q$ , onde cada conjunto  $G_i$  contém as localidades em que uma das  $q$  variantes selecionadas ocorrem,  
 Seja  $S$  um conjunto de conjuntos de localidades,  $S \leftarrow \emptyset$ ,

Seja  $T$  um conjunto de localidades,

Se  $q = 4$ ,

Então:

$T \leftarrow G_1 \cap G_2 \cap G_3 \cap G_4$

Se  $T \neq \emptyset$ ,

Então:

$S = \{T\}$

Se  $q \geq 3$ ,

Então:

Para todo  $G_i$  pertencente a  $G$ ,

Para todo  $G_j$  pertencente a  $G$ ,  $G_j \neq G_i$ ,

Para todo  $G_k$  pertencente a  $G$ ,  $G_k \neq$

$G_i, G_k \neq G_j$ ,

$T \leftarrow G_i \cap G_j \cap G_k$

Se  $T \neq \emptyset$  E  $(T \neq S_i)$ , para

todo  $S_i$  pertencente a  $S$ )

Então:

$S = S \cup \{T\}$

Se  $q \geq 2$ ,

Então:

Para todo  $G_i$  pertencente a  $G$ ,

Para todo  $G_j$  pertencente a  $G$ ,  $G_j \neq G_i$ ,

$T \leftarrow G_i \cap G_j$

Se  $T \neq \emptyset$  e  $(T \neq S_i)$ , para todo  $S_i$

pertencente a  $S$ )

Então:

$S = S \cup \{T\}$

Se  $q \geq 1$ ,

Então,

Para todo  $G_i$  pertencente a  $G$ ,

$T \leftarrow G_i$

Se  $T \neq \emptyset$  E  $(T \neq S_i)$ , para todo  $S_i$  pertencente

a  $S$ )

Então:

$S = S \cup \{T\}$

Fig. 4. Algoritmo Cálculo das Interseções nas Isoglossas

Em seguida, as áreas correspondentes são coloridas com as cores geradas e exibidas sobre o mapa, obtendo o resultado ilustrado na Fig. 5.

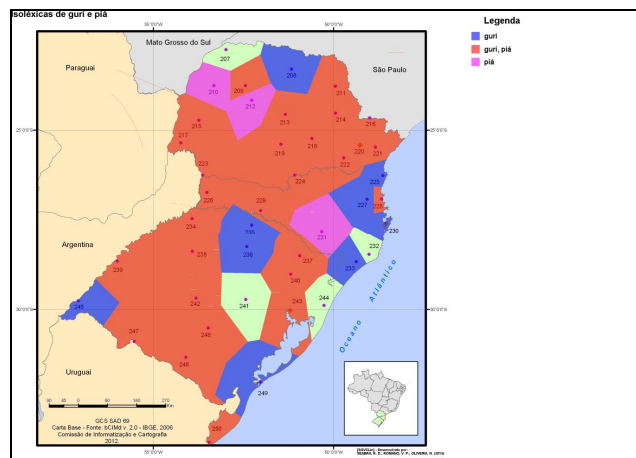


Fig. 5. Exemplo de carta isoglossa para as variantes *guri* e *piá*. As áreas indicadas em azul apontam a ocorrência de *guri*; em magenta, de *piá*; e, em vermelho, as interseções de ambas

**C. Cálculo dos Gradientes nas Isoglossas**

O cálculo dos gradientes nas isoglossas é complexo, envolvendo a geração de cores em tempo de execução, além do cálculo de grandes volumes de dados. Esse tipo de cálculo, em particular, só pode ser

aplicado a uma variante por vez. Os Algoritmos 3 e 4 (Fig. 5 e Fig. 6, respectivamente) apresentam o modo de execução do cálculo envolvido nessa funcionalidade.

**Algoritmo 3: Cálculo das Cores dos Gradientes**

Seja  $cria\_cor$  ( $i$  um inteiro,  $max$  um inteiro) uma função definida pelo seguinte algoritmo:

---

```

Seja  $H \leftarrow 5 / 9 + i / (9 * \text{max})$ 
Seja  $S \leftarrow 1 / 2 + i / (2 * \text{max})$ 
Seja  $V \leftarrow 1 - i / (2 * \text{max})$ 
Seja R, G, B os valores de cada cor a serem calculados
Seja A  $\leftarrow 200$ , o valor de opacidade da cor
Se S = 0,
Então:
    R  $\leftarrow V * 255$ ; G  $\leftarrow R$ ; B  $\leftarrow R$ 
Senão:
    Seja h  $\leftarrow (H - \lfloor H \rfloor) * 6$ 
    Seja f  $\leftarrow H - \lfloor H \rfloor$ 
    Seja p  $\leftarrow V * (1 - S)$ 
    Seja q  $\leftarrow V * (1 - S * f)$ 
    Seja t  $\leftarrow V * (1 - (S * (1 - f)))$ 
    Se  $\lfloor H \rfloor = 0$ ,
    Então:
        R  $\leftarrow \lfloor V * 255 + 0.5 \rfloor$ ; G  $\leftarrow \lfloor t * 255 + 0.5 \rfloor$ ;
        B  $\leftarrow \lfloor p * 255 + 0.5 \rfloor$ 
    Senão Se  $\lfloor H \rfloor = 1$ ,
    Então:
        R  $\leftarrow \lfloor q * 255 + 0.5 \rfloor$ ; G  $\leftarrow \lfloor V * 255 + 0.5 \rfloor$ ;
        B  $\leftarrow \lfloor p * 255 + 0.5 \rfloor$ 
    Senão Se  $\lfloor H \rfloor = 2$ ,
    Então:
        R  $\leftarrow \lfloor p * 255 + 0.5 \rfloor$ ; G  $\leftarrow \lfloor V * 255 + 0.5 \rfloor$ ;
        B  $\leftarrow \lfloor t * 255 + 0.5 \rfloor$ 
    Senão Se  $\lfloor H \rfloor = 3$ ,
    Então:
        R  $\leftarrow \lfloor p * 255 + 0.5 \rfloor$ ; G  $\leftarrow \lfloor q * 255 + 0.5 \rfloor$ ;
        B  $\leftarrow \lfloor V * 255 + 0.5 \rfloor$ 
    Senão Se  $\lfloor H \rfloor = 4$ ,
    Então:
        R  $\leftarrow \lfloor t * 255 + 0.5 \rfloor$ ; G  $\leftarrow \lfloor p * 255 + 0.5 \rfloor$ ;
        B  $\leftarrow \lfloor V * 255 + 0.5 \rfloor$ 
    Senão Se  $\lfloor H \rfloor = 5$ ,
    Então:
        R  $\leftarrow \lfloor V * 255 + 0.5 \rfloor$ ; G  $\leftarrow \lfloor p * 255 + 0.5 \rfloor$ ;
        B  $\leftarrow \lfloor q * 255 + 0.5 \rfloor$ 
Seja cor  $\leftarrow A \ll 24 \mid R \ll 16 \mid G \ll 8 \mid B$ 
Retorna cor

```

---

Fig. 7. Algoritmo Cálculo das Cores dos Gradientes

---

#### Algoritmo 4: Cálculo dos Gradientes

---

```

Seja inteiro (valor) uma função: inteiro(valor)  $\mapsto \lfloor \text{valor} + 0.5 \rfloor$ 
Seja distancia  $(x_1, y_1, x_2, y_2)$  uma função: distancia  $(x_1, y_1, x_2, y_2) \mapsto \text{sqrt}((x_1 - x_2)^2 + (y_1 - y_2)^2)$ 
Seja I uma matriz de pontos (x,y) contendo todos os pontos do mapa a ser gerado,
Seja P um conjunto de pontos valorados contendo todas as localidades e suas respectivas contagens da variante a ser trabalhada,
Seja max  $\leftarrow 0$ 
Para todo  $P_i$  pertencente a P,
    Se  $P_i.\text{valor} > \text{max}$ ,
    Então:
        max  $\leftarrow P_i.\text{valor}$ 
Seja C um array de cores,  $|C| = \text{max} + 1$ 
 $C_1 \leftarrow$  cor transparente
Para todo  $C_i$  pertencente a C,  $i \neq 1$ 
     $C_i \leftarrow$  cria_cor(i, max)
Para todo  $I_{ij}$  pertencente a I,
    Seja soma  $\leftarrow 0$ 
    Seja valor  $\leftarrow 0$ 
    Para todo  $P_k$  pertencente a P,
        Seja d  $\leftarrow$  distancia( $i, j, P_k.x, P_k.y$ )  $^4$ 
        soma  $\leftarrow$  soma + d
        valor  $\leftarrow$  valor + d *  $P_k.\text{valor}$ 
     $I_{ij} \leftarrow C_{\text{inteiro}(\text{valor} / \text{soma})}$ 

```

---

Fig. 8. Algoritmo Cálculo dos Gradientes

Com esse processamento, é gerado o efeito de gradiente ilustrado na Fig. 7.

A Fig. 8 ilustra um exemplo de mapeamento dos gradientes a partir de uma visualização tridimensional de parte dos pontos valorados. Pode-se notar a continuidade, de um ponto a outro, calculada pelo algoritmo e, também, a separação de cada curva de nível em uma tonalidade da cor azul.



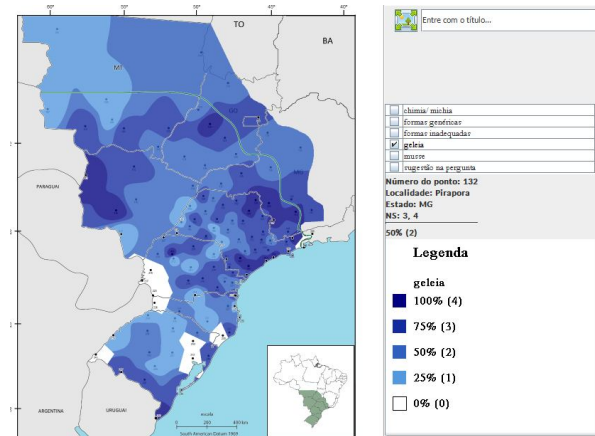


Fig. 9. Exemplo de carta de isoglossa com uso de gradientes para a variante *geleia*. O uso de diferentes tons indica as variações, em função da produtividade, das ocorrências da variante linguística nas localidades do mapa.

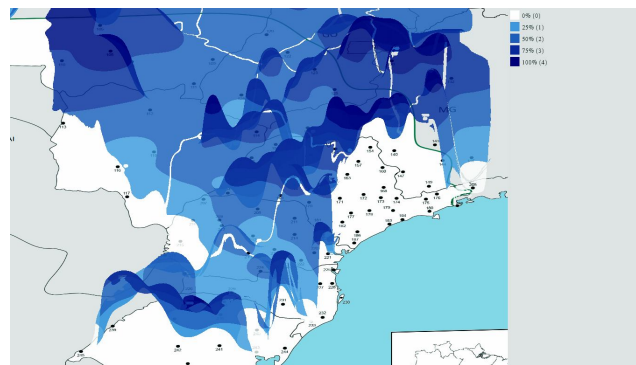


Fig. 10. Visualização tridimensional dos gradientes. Os dados da interpolação podem ser representados no mapa como uma superfície de nível, de forma a facilitar a visualização das faixas de valores e suas respectivas tonalidades.

Analisando a complexidade desse algoritmo, sendo  $m$  a quantidade de linhas da imagem,  $n$  a quantidade de colunas e  $p$  a quantidade de pontos valorados a serem processados, tem-se que a complexidade é  $O(m * n * p)$ . Face a isso, e também pelo fato de que cada ponto da imagem pode ser calculado de forma independente em relação aos outros, o algoritmo foi paralelizado. Assim, cada linha da imagem foi calculada em uma *thread*, e o resultado do cálculo armazenado na imagem. Com esse recurso, foi obtido um ganho significativo de desempenho, conforme mostra a Fig. 9.

Os testes foram realizados em um notebook com um processador Intel Core i3 M370 (dois núcleos e quatro *threads* a 2.4GHz), e com 4GB de memória RAM. Para o cálculo foram criados 50 pontos valorados e espalhados aleatoriamente sobre uma imagem quadrada com 5000 *pixels* de lado. Para esses pontos, foram atribuídos valores também aleatórios.

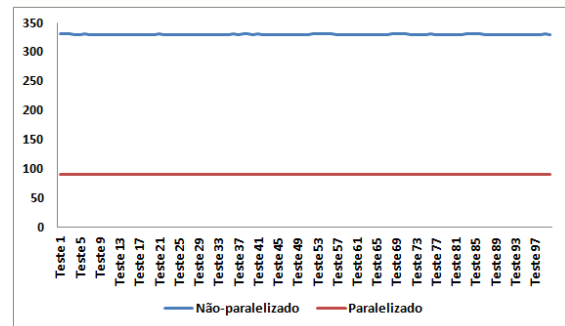


Fig. 11. Tempo de execução do algoritmo em segundos. A aplicação da paralelização mostra resultados melhores.

#### D. Cálculo das Cartas Diatópicas

Nas cartas diatópicas – que representam as ocorrências de uma ou mais variantes em uma rede de pontos linguísticos – foi necessária a exibição de gráficos em formato pizza contendo as quantidades referentes a cada um dos pontos linguísticos. Ademais, é exibido um histograma gerado a partir da soma dos dados linguísticos de todos os pontos. Para a geração desses gráficos, os dados são calculados e enviados a

uma biblioteca de geração de gráficos – *JfreeChart* [39]. O resultado que caracteriza um exemplo de carta

diatópica pode ser conferido na Fig. 10.

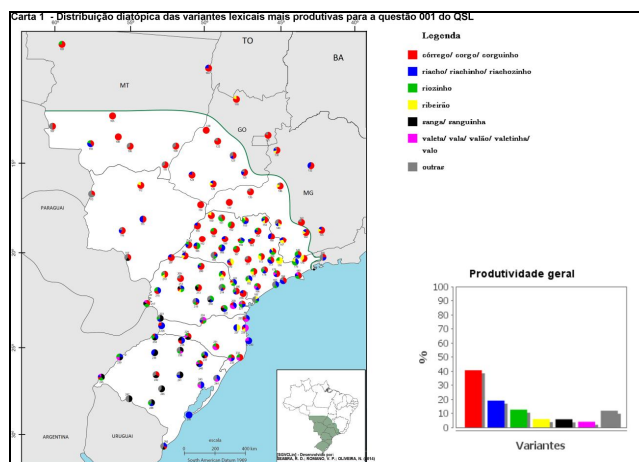


Fig. 12. Exemplo de carta diatópica para a variante *córrego*. As ocorrências das variações lexicais são exibidas, em cada ponto linguístico, em gráficos com formato pizza.

## V. CONCLUSÃO

A necessidade de otimização do tempo e a participação direta do linguista na elaboração de cartas linguísticas faz do [SGVCLin] uma ferramenta inovadora, pois, atualmente, não existe um programa computacional desenvolvido de forma que possa ser amplamente utilizado em trabalhos geolinguísticos diversos. Esta flexibilidade atribuída à ferramenta deve-se ao fato de que ela foi desenvolvida considerando-se muitos dos aspectos metodológicos empregados no Projeto ALiB, referência em Geolinguística brasileira.

O usuário da ferramenta, no caso o linguista, consegue interagir naturalmente com as funcionalidades do sistema, uma vez que tem à sua disposição um tutorial explicativo. Ademais, os botões e ícones, bem como as *tooltips* disponíveis na interface indicam os passos executados em cada etapa, propiciando uma interação intuitiva e didática.

É de suma importância a divulgação de pesquisas que utilizam os avanços tecnológicos disponíveis para solucionar problemas metodológicos de outras áreas. Nesse sentido, o [SGVCLin] não apenas agiliza o processo de cartografia, como também fornece relatórios diversos que cruzam informações linguísticas e extralinguísticas, avançando aos “veios sociolinguísticos” [17] da dialetologia contemporânea.

O trabalho apresentado neste artigo é parte desta ferramenta inovadora para a geração e a visualização de cartas linguísticas. Devido à necessidade de desenvolvimento de funcionalidades específicas ao software implementado, neste trabalho, foram apresentadas algumas soluções na forma de algoritmos concebidos para três aspectos principais: (i) cálculo da área de preenchimento de cartas isoglossas; (ii) cálculo das interseções nas isoglossas; e (iii) cálculo para a geração das cores e gradientes nas isoglossas. Testes realizados mostraram que o uso de implementação

paralela do algoritmo responsável pela geração dos gradientes nas isoglossas acelerou significativamente o desempenho do algoritmo em questão.

Um recurso não obrigatório para o funcionamento da ferramenta mas que pode privilegiar ainda mais a visualização dos dados linguísticos apresentados nas cartas diatópicas e de isoglossas, os trabalhos futuros envolverão a opção de o usuário interagir com as cartas por meio de recursos que extrapolem a visualização bidimensional. Visando a alcançar esse objetivo, serão utilizados recursos de Realidade Aumentada [40, 41] na visualização dos gráficos exibidos nas cartas linguísticas.

O trabalho realizado não se apresenta como definitivo, pelo contrário, levanta novas questões de investigação. Partindo-se dessa premissa, os autores ainda realizarão um estudo acerca da avaliação da ferramenta, sob o ponto de vista de professores, estudantes e pesquisadores da área da geolinguística, segundo critérios de usabilidade.

## REFERÊNCIAS

- [1] J. Gilliéron and E. Edmont, “Atlas Linguistique de la France”. 35 fac. Paris: Honoré Champion, 1902-1910, 1915.
- [2] H. Kurath et al., “Linguistic Atlas of New England (LANE)”. 3 vol. Brown University Press, 1939-1943.
- [3] T. A. Navarro, “El español en Puerto Rico: Contribuciones a la geografía lingüística hispanoamericana”, 2. ed. Porto Rico: Rio Piedras, Editorial Universitaria, Universidad de Puerto Rico, 1966 [1946].
- [4] V. P. Romano, Balanço crítico da Geolinguística brasileira e a proposição de uma divisão. *Entretextos*. Londrina, UEL, v.13, n.2, p. 203-242, 2013.
- [5] N. Rossi et al., “Atlas Prévio dos Falares Baianos”. Rio de Janeiro: INL, 1963.
- [6] V. P. Romano, R. D. Seabra e N. Oliveira, “[SGVCLin] – Software para geração e visualização de cartas linguísticas”, *Revista de Estudos da Linguagem*, v.22, p.119-151, 2014.
- [7] J. Canny, “A computational approach to edge detection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v.8, n.6, pp. 679-698, 1986.

- [8] D. Shepard, "A two-dimensional interpolation function for irregularly-spaced data", *ACM '68 Proceedings of the 1968 23rd ACM National Conference*, New York, pp.517-524, 1968.
- [9] A. Houaiss, M. Villar, "Dicionário eletrônico Houaiss da língua portuguesa". Versão 1.0, 2001.
- [10] E. Coseriu, "O homem e a sua linguagem". Tradução de Carlos Alberto da Fonseca e Mário Ferreira. Rio de Janeiro: Presença, 1987.
- [11] S. F. Brandão, "A geografia linguística no Brasil". São Paulo: Ática, 1991.
- [12] C. Ferreira, "Atlas Prévio dos Falares Baianos: alguns aspectos metodológicos". In.: V. Aguilera. Geolinguística no Brasil – caminhos e perspectivas. Londrina: EDUEL, pp.15-29, 1998.
- [13] J. Ribeiro et al., "Esboço de um atlas linguístico de Minas Gerais". Rio de Janeiro: Casa de Rui Barbosa, 1977.
- [14] M. do S. S. de Aragão e C. Bezerra de Menezes, "Atlas linguístico da Paraíba". Brasília: UFPB; CNPq, Coordenação Editorial, 1984.
- [15] C. Ferreira et al., "Atlas linguístico de Sergipe", Salvador: Universidade Federal da Bahia, Fundação de Cultura de Sergipe, 1987.
- [16] V. de Aguilera, "Atlas linguístico do Paraná", Curitiba: Imprensa Oficial, 1994.
- [17] S. A. M. Cardoso, "Geolinguística: tradição e modernidade", São Paulo: Parábola, 2010.
- [18] A. Razky, "Atlas linguístico sonoro do Estado do Pará (ALiSPA I.1)". Belém: s/ed. 2004 (programa em CD-ROM).
- [19] L. G. Freitas Marins, "O rural e o urbano: novos e velhos olhares no Centro-Oeste do Brasil", 2012. 325f. Dissertação (Mestrado em Estudos de Linguagens) – Universidade Federal do Mato Grosso do Sul, Campo Grande, 2012.
- [20] V. C. M. Benke, "Tabus linguísticos nas capitais do Brasil: um estudo baseado em dados geolinguísticos", 2012. 310f. Dissertação (Mestrado em Estudos de Linguagens) – Universidade Federal do Mato Grosso do Sul, Campo Grande, 2012.
- [21] D. de S. S. Costa, "Léxico indígena nas capitais brasileiras: um estudo geolinguístico", 2012. 209f. Dissertação (Mestrado em Estudos de Linguagens) – Universidade Federal do Mato Grosso do Sul, Campo Grande, 2012.
- [22] H. G. Vieira, "Fundamentos para organizar, implementar e manter um banco de dados geolinguístico". In.: V. de A. Aguilera, A geolinguística no Brasil: caminhos e perspectivas. Londrina: EDUEL, pp.207-224, 1998.
- [23] W. Koch, C. Altenhofen e M. S. Klassman (orgs.), "Atlas Linguístico-Emográfico da Região Sul do Brasil – ALERS": cartas fonéticas e morfosintáticas, 1. ed., Porto Alegre: Ed. UFRGS, Florianópolis: Ed. UFSC, 2002.
- [24] D. G. Oliveira (org.), "Atlas linguístico do Mato Grosso do Sul – ALMS", Campo Grande: Ed. UFMS, 2007.
- [25] M. L. de C. Cruz, "Atlas linguístico do Amazonas", 2004. 2v. Tese (Doutorado em Língua Portuguesa), Universidade Federal do Rio de Janeiro, Rio de Janeiro, 2004.
- [26] A. Conci, E. Azevedo e F. R. Leta, "Computação gráfica", v.2. Rio de Janeiro: Elsevier, 2008.
- [27] H. Pedrini e W. R. Schwartz, "Análise de imagens digitais: Princípios, algoritmos e aplicações". São Paulo: Thomson Learning, 2008.
- [28] F. Aurenhammer, "Voronoi diagrams – a survey of a fundamental geometric data structure", *ACM Computing Surveys*, New York, v. 23, n. 3, pp.345-405, 1991.
- [29] A. Okabe et al., "Spatial tessellations: concepts and applications of Voronoi diagrams". New York: John Wiley & Sons, 2009.
- [30] P. Zamperoni, "Analysis of some region growing operators for image segmentation", *International Conference on Advances in Image and Processing Pattern Recognition*, Amsterdam, Holanda, pp. 204-208, 1986.
- [31] W. Grimson and T. Lozano-Perez, "Localizing overlapping parts by searching the interpretation tree", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v.9, n.4, pp. 469-482, 1987.
- [32] R. M. Haralick and L. G. Shapiro, "Computer and Robot Vision", v.1, Addison-Wesley, Reading, MA, Estados Unidos, 1992.
- [33] D. Maar and E. C. Hildreth, "Theory of edge detection", *Proceedings of the Royal Society of London B207*, pp. 187-217, 1980.
- [34] V. Berzins, "Accuracy of Laplacian edge detection", *Computer Vision, Graphics, Image Processing*, v. 27, n. 2, pp. 195-210, 1984.
- [35] A. Huertas and G. Medioni, "Detection of intensity changes with subpixel accuracy using Laplacian-Gaussian masks", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 8, n. 5, pp. 651-664, 1986.
- [36] T. Lindeberg, "Edge detection and ridge detection with automatic scale selection", *International Journal of Computer Vision*, v. 30, n. 2, pp. 117-154, 1998.
- [37] S. A. Cardoso e J. A. Mota, "Projeto Atlas Linguístico do Brasil: antecedentes e estágio atual", *Alfa: Revista de Linguística*, São Paulo, v. 56, n. 3, pp.855-870, 2012.
- [38] H. Thun, "La géographie linguistique romane à la fin du XX siècle". In.: D. V. Raendonck et al. (Orgs), *Actes du XXII Congrès International de Linguistique e Philologie Romanes*. Bruxelles, 1998, p. 367-409.
- [39] Object Refinery Limited, "JfreeChart", Disponível em: <<http://www.jfree.org/jfreechart/>>. Acesso em: 22 de março de 2014.
- [40] R. Tori, C. Kirner e R. Siscoutto, "Fundamentos e tecnologia de realidade virtual e aumentada", Belém: Editora SBC, 2006, pp.20-34.
- [41] C. Kirner e R. Santin, "Interaction, collaboration and authoring in augmented reality environments", *Anais do XI Symposium on Virtual and Augmented Reality*, Porto Alegre: SBC, 2009, pp.210-220.