

UFPB – UNIVERSIDADE FEDERAL DA PARAÍBA
BACHARELADO EM MATEMÁTICA COMPUTACIONAL

DIEGO ALBERTIM MEIRA

**ANÁLISE DA AMOSTRAGEM POR PONTO MAIS DISTANTE PARA A
APROXIMAÇÃO GEODÉSICA EM UM GRAFO: O ESTADO DA ARTE**

JOÃO PESSOA-PB

2017

DIEGO ALBERTIM MEIRA

**ANÁLISE DA AMOSTRAGEM POR PONTO MAIS DISTANTE PARA A
APROXIMAÇÃO GEODÉSICA EM UM GRAFO: O ESTADO DA ARTE**

Monografia apresentada à Universidade Federal da Paraíba – UFPB como requisito parcial para obtenção do título de Bacharel em Matemática Computacional sob orientação do Prof. Dr. Sérgio de Carvalho Bezerra

JOÃO PESSOA – PB

2017

Ficha Catalográfica elaborada por
Rogério Ferreira Marques CRB15/690

M499a Meira, Diego Albertim.
Análise da amostragem por ponto mais distante para a aproximação
geodésica em um grafo: o estado da arte / Diego Albertim Meira. – João
Pessoa, 2017.
34p. : il.

Monografia (Bacharelado em Matemática Computacional) –
Universidade Federal da Paraíba - UFPB.
Orientador: Profº. Dr. Sergio de Carvalho Bezerra.

1. Algoritmo. 2. Grafo. 3. Geodésica. 4. Matemática computacional. I.
Título.

UFPB/BSCI

CDU: 004.021 (043.2)



CENTRO DE INFORMÁTICA
UNIVERSIDADE FEDERAL DA PARAÍBA

ATA DE DEFESA PÚBLICA DO TRABALHO DE CONCLUSÃO DE CURSO

Aos 14 dias do mês de junho de 2017, às 14:00 horas, em sessão pública na sala T- 01 do Campus I da Universidade Federal da Paraíba, na presença da banca examinadora presidida pelo professor(a) orientador Sérgio de Carvalho Bezerra e pelos professores Andréa Vanessa Rocha e Alexandre Simas Bustamantes, o aluno Diego Albertim Meira, apresentou o trabalho de conclusão de curso intitulado: ANÁLISE DA AMOSTRAGEM POR PONTO MAIS DISTANTE PARA A APROXIMAÇÃO GEODÉSICA EM UM GRAFO: O ESTADO DA ARTE como requisito curricular indispensável para a integralização do Curso de Matemática Computacional.

Após a exposição oral, o(a) candidato(a) foi arguido(a) pelos componentes da banca que reuniram-se reservadamente, e decidiram, *aprovar* a monografia, com nota 9,5. Divulgando o resultado formalmente ao aluno e demais presentes, eu, na qualidade de Presidente da Banca, lavrei a presente ata que será assinada por mim, pelos demais examinadores e pelo aluno.

Sérgio de Carvalho Bezerra

<Presidente da Banca Examinadora>

Andréa Vanessa Rocha

<Examinador 01>

Alexandre de B. Simas

<Examinador-02>

Diego Albertim Meira

<Aluno>

AGRADECIMENTOS

Gostaria de agradecer a minha família, amigos, colegas e professores que me incentivaram a seguir em frente no curso e me apoiaram nos momentos de dúvida que surgiram durante minha permanência no mesmo. Em especial, gostaria de agradecer a meu orientador, prof. Sérgio de Carvalho, por sua paciência para comigo e sua orientação neste trabalho, à prof.^a Andrea Rocha, por me indicá-lo e por ter me ajudado a clarear algumas dúvidas que me surgiram no decorrer do curso, ao prof. André Torii, por me apresentar a área que possivelmente irei estudar em uma pós-graduação, e a meu amigo e antigo colega de curso Pablo Luna, por me ajudar a desvendar a formatação que precisei aplicar no texto.

*"Apegue-se a seus sonhos, pois seus
pesadelos podem parecer que são a sua
realidade"*

(Jacoby Shaddix)

Resumo

O algoritmo de amostragem por ponto mais distante, ou *farthest point sampling*, tem sido amplamente utilizado em diversas aplicações que envolvam processamentos de imagem, mapeamentos de superfícies, entre outras finalidades. Um recente trabalho, proposto pela equipe do prof. Sylvain Lazard, aplica este algoritmo para aproximar distâncias geodésicas em grafos associados a superfícies isometricamente invariantes, sendo que alguns pontos na publicação se mostraram um pouco nebulosos. Neste presente trabalho, será oferecida uma complementação bibliográfica ao projeto de Lazard et al, no intuito de clarificar questões acerca do método de *farthest point sampling* e alguns outros conceitos utilizados na aplicação proposta. Além disso, tenciona-se oferecer uma tradução satisfatória das ideias do referido trabalho.

Palavras-chave: amostragem por ponto mais distante, algoritmo, geodésicas, grafo, complementação.

Abstract

The farthest point sampling algorithm has been vastly used in various applications involving image processing, surface mapping, among other purposes. A recent work, proposed by the team of prof. Sylvain Lazard, applies this algorithm to approximate geodesic distances in graphs associated to isometry-invariant surfaces, being that some points in the publishing showed to be a little nebulous. In this present work, will be offered a bibliographic complementation to Lazard et al's project, intending to clarify some questions about the farthest point sampling method and other concepts used on the proposed application. Moreover, it's intended to offer a satisfactory translation of the ideas in the referred work.

Keywords: farthest point sampling, algorithm, geodesics, graph, complementation.

LISTA DE FIGURAS

Figura 1	12
Figura 2	14
Figura 3	27

LISTA DE ABREVIATURAS

APD – Amostragem por Ponto mais Distante (Farthest Point Sampling)

DV – Diagrama de Voronoi

TD – Triangulação de Delaunay

CMOU – Caminho Mínimo de Origem Única (Single Source Shortest Path)

CMTD – Caminho Mínimo de Todos os Pares (All Pairs Shortest Path)

DVL – Diagrama de Voronoi Limitado

SUMÁRIO

1 Introdução	11
2 O Algoritmo APD	13
2.1 A estratégia de ponto mais distante	13
2.2 Diagramas de Voronoi	13
2.2.1 Triangulação de Delaunay	15
2.2.2 Diagrama de Voronoi Limitado	16
2.3 Algoritmo	17
3 Trabalhos relacionados e adaptações ao algoritmo	18
4 O modelo para aproximação geodésica em um grafo	19
4.1 Definição do problema	20
4.2 O uso do algoritmo APD	21
4.3 Fator de alongamento	22
4.4 Complexidade do método	27
5 Conclusões	32
Referências Bibliográficas	33

1. Introdução

Na computação atual, diversas tarefas, como modelagens de superfícies, ou a transferência de uma imagem, requerem o cálculo da distância entre pontos em superfícies não-euclidianas, sendo que esta distância é aproximada como uma curva, recebendo a denominação de distância geodésica [9].

Tais tarefas costumam ser tratadas através de algoritmos que realizam amostragem progressiva para calcular a distância entre pontos. Entretanto, como notado em [2], o uso destes algoritmos pode gerar problemas, tais como falhas de transmissão e surgimento de ruído para, por exemplo, a transmissão de imagens de alta resolução. Além disso, tais métodos não permitem que novos pontos sejam incluídos na amostragem com facilidade.

Tendo isto em mente, Eldar et al, propuseram em [2] um novo método algorítmico para resolver tais problemas com as amostras, chamando o método de “amostragem por ponto mais distante” (farthest point sampling, ou APD). Este modelo inovador acabou se provando eficaz e passou a ser utilizados em tarefas que exigiam a manipulação de formas irregulares, chegando inclusive a receber, com o passar dos anos, algumas adaptações propostas por diferentes pesquisadores, para a resolução de problemas específicos.

Dentre estes pesquisadores, ao final de 2016, Lazard et al utilizaram o método em [1] numa forma de associar as superfícies a serem abordadas com grafos, provando que a distância entre os vértices desses grafos é no máximo a distância geodésica entre os pontos. Os resultados obtidos foram utilizados na comprovação da eficácia do APD para processamento de formas isometricamente invariantes (ou seja, cujas distâncias são preservadas mediante transformações geométricas).

Entretanto, ao avaliar o artigo proposto por esses pesquisadores, verifica-se uma certa deficiência didática no que diz respeito ao próprio algoritmo APD e alguns conceitos utilizados na obtenção dos resultados apresentados. Nesse sentido, este presente trabalho busca complementar tais pontos, no intuito de melhorar a

compreensão do artigo original, bem como oferecer uma tradução satisfatória do mesmo para a língua portuguesa.

Neste trabalho, será apresentado, na seção 2, os conceitos utilizados na construção do método e o algoritmo propriamente dito. Na seção 3, serão mostrados trabalhos relativos ao uso do APD e outros que apresentaram adaptações ao método original. Na seção 4, serão apresentados os teoremas que compõem o trabalho proposto por Lazard et al (sendo estes o foco principal deste trabalho). Por fim, será dada uma rápida conclusão sobre o assunto.

2. O algoritmo APD

O APD é, conforme [2], um método para gerar uma sequência de pontos planares com distribuição espacial uniforme para qualquer número de amostras. Este algoritmo baseia-se no conceito de amostragem progressiva, funcionando como uma alternativa aos modelos estocásticos baseados na distribuição de Poisson. Para o seu desenvolvimento, dois pontos principais foram levantados; a estratégia para desenvolvê-lo e uma forma eficiente de implementação. Tais tópicos serão abordados a seguir.

2.1. A estratégia de ponto mais distante

O método segue uma estratégia denominada de Estratégia de Ponto mais Distante, onde dado um ponto em uma região amostral, o próximo ponto amostral será aquele que se encontra mais distante das demais amostras anteriores (geometricamente, o ponto que representa o centro de uma circunferência localizada dentro dos limites da região). Formalmente, esta estratégia é definida como:

Definição 1: Dado um conjunto de pontos amostrais $S = s_i, 0 < i < n - 1$, com n pontos, definido em uma região A (por exemplo, um retângulo), o próximo ponto amostral será o ponto p , de forma que:

$$p = \operatorname{argmax}_{q \in A} d(q, S), \text{ onde } d(q, S) = \max_{q \in A} \left(\min_{0 \leq i < n-1} d(q, s_i) \right)$$

2.2. Diagramas de Voronoi

Os pesquisadores perceberam que a implementação de tal estratégia poderia ser melhor realizada através da construção de diagramas de Voronoi, estruturas de uso bastante difundido na computação, para a resolução de problemas de

proximidade. As unidades básicas desse diagrama, ditas células de Voronoi, são definidas da seguinte forma:

Definição 2: Seja $S = s_i, 0 < i < n - 1$ um conjunto planar de pontos. A célula de Voronoi do ponto s_i , $V(s_i)$, é definida como o conjunto de todos os pontos planares mais próximos de s_i do que qualquer outro ponto de S . A célula é um polígono convexo, mas não necessariamente limitado, com não mais que $n - 1$ lados.

O particionamento do plano, onde S é definido, em células de Voronoi é chamado Diagrama de Voronoi de S , e denotado por $DV(S)$. Os limites das células constituem as arestas do diagrama, e suas interligações, os vértices. Além disso, caso dois pontos $s_i, s_j \in S$ estejam em células que compartilhem uma aresta em comum, estes serão ditos pontos vizinhos.

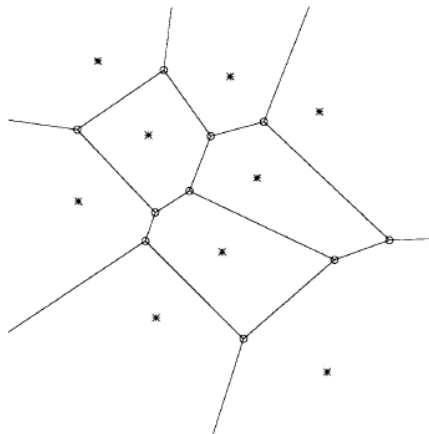


Figura 1: Diagrama de Voronoi para um conjunto amostral S

Para prosseguir na construção do algoritmo, os pesquisadores adaptaram um pequeno grupo de propriedades inerentes à construção dos DV's, assumindo que não mais que três pontos s_i pertençam à mesma circunferência que geometricamente define as posições pela estratégia APD. Tais propriedades são:

- 1) Todo vértice de Voronoi é a intersecção de exatamente três arestas, implicando que todo vértice $v \in DV(S)$ é o centro de uma circunferência $C(v)$ que passa por três pontos de S ;
- 2) Para cada vértice de Voronoi $v \in DV(S)$, a circunferência $C(v)$ não contém nenhum ponto de S além dos três que já atravessa;
- 3) Uma célula de Voronoi de um ponto $s \in S$ não é limitada se, e somente se, o ponto s está posicionado na borda do envoltório convexo de S .

2.2.1. Triangulação de Delaunay

Para além do uso dos diagramas de Voronoi, adicionalmente foi visitado o conceito da triangulação de Delaunay. Uma triangulação é definida como sendo um conjunto de segmentos de linha, que não se cruzam, ligando os pontos de um conjunto de pontos planares S , e dividindo seu envoltório convexo em triângulos.

Considerando uma triangulação T de S , esta será dita triangulação de Delaunay (denotada por $TD(S)$), se para cada triângulo em T a circunferência que o circunscreve não contém nenhum ponto de S além dos três que formam o triângulo. Ainda mais, essa triangulação é única e ótima de um ponto de vista angular (desconsiderando casos bizarros, como superfícies muito irregulares).

Uma dualidade entre o processo de TD e os DV's pode ser verificada pela segunda propriedade descrita na subseção anterior, onde cada vértice de Voronoi corresponderá a um triângulo na TD. Assim, a construção de $DV(S)$ provém uma triangulação ótima para S .

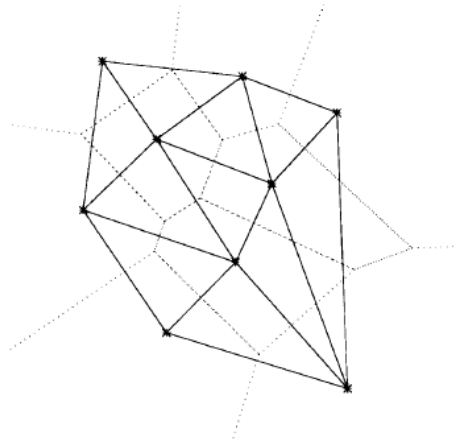


Figura 2: Triangulação de Delaunay sobre o Diagrama de Voronoi, onde os vértices dos triângulos representam os pontos amostrais

2.2.2. Diagrama de Voronoi Limitado

Como verificado anteriormente, o APD requer que, a cada estágio da amostragem, seja encontrado o ponto que está mais distante do conjunto atual de amostras S , e que tal é realizado através da construção de diagramas de Voronoi, que produz uma distribuição espacial para as amostras de S . Sendo o domínio trabalhado finito, considera-se uma célula de Voronoi limitada associada a cada ponto amostral. Tal pode ser definido, usando-se o conceito de processamento de imagens, como a conjunção entre a célula de Voronoi e a área de definição da imagem.

A união entre todas as células associadas ao conjunto S fornece o diagrama de Voronoi limitado $DVL(S) = (V, E)$, onde V e E representam, respectivamente, o conjunto de vértices e o conjunto de arestas do diagrama. Assumindo que, por exemplo, se trabalhe com uma imagem retangular, onde as amostras são localizadas em seus cantos, tem-se que cada célula de Voronoi limitada conterá o ponto amostral, e todos os pontos da imagem mais próximos a ele que qualquer outro ponto amostral. Isso implica, que o ponto mais distante do atual está localizado num dos vértices do DVL.

2.3. Algoritmo

Tendo em mente esses conceitos e os apresentados nas subseções anteriores, o algoritmo APD pode enfim ser descrito da seguinte forma:

(1) Crie um conjunto de pontos iniciais S^n consistindo de todos os cantos da imagem, e um ponto adicional escolhido aleatoriamente.

Calcule $DVL(S^n) = (V^n, E^n)$.

(2) Em cada célula i encontre o ponto $p \in V^n$ que esteja mais distante de s_i
Escolha o ponto p que maximiza essa distância

(3) Faça $S^{n+1} = S^n \cup \{p\}$

Calcule $DVL(S^{n+1})$

Faça $n = n + 1$

(4) Caso mais amostras sejam necessárias, retorne para (2).

O critério de parada exato para este algoritmo dependerá da aplicação ao qual ele será sujeito, podendo, teoricamente, o processo de amostragem seguir infinitamente.

3. Trabalhos relacionados e adaptações ao algoritmo

Após sua publicação original, o uso do algoritmo APD começou a ganhar certa popularidade entre estudiosos que trabalhavam com mapeamento de superfícies e processamento de imagens, por suas propriedades antirruído e sua alta taxa de aquisição de dados [11]. Além disso, versões adaptadas do algoritmo para tarefas específicas também foram propostas, mostrando o crescente interesse no uso deste método.

Em [11], o algoritmo foi adaptado para utilizar-se da técnica de *Fast Marching*, utilizada para o estudo de problemas de propagação em domínios limitados, acrescentando eficiência ao método e permitindo a extensão de seu uso para casos de amostragem não-uniforme e para domínios com pesos diferentes associados a pontos diferentes na superfície.

Em [10], a técnica citada anteriormente foi incrementada ao utilizar-se uma extensão da técnica de *Fast Marching*, permitindo que o APD pudesse ser usado em problemas envolvendo superfícies implícitas e nuvens de pontos. Pouco tempo depois, esta adaptação foi usada em [7] para a construção de um algoritmo para simplificação de nuvens de pontos, permitindo processamentos mais eficientes nesse tipo de superfície.

Algumas outras aplicações do algoritmo podem ser vistas em [6], onde o mesmo foi utilizado na aceleração do processamento de malhas espectrais, em [15] e [16], tendo sido aplicado na construção de algoritmos de pareamento de superfícies isometricamente invariantes, em [17], onde o modelo foi adaptado para a compressão de imagens trianguladas de forma anisotrópica (alguma propriedade física varia com mudanças na sua direção), e em [4], onde foi usado para o desenvolvimento de um método de registro facial em três dimensões baseado em modelos de deformação e curvas spline.

4. O modelo para aproximação geodésica em um grafo

Em seu trabalho de 2016, Lazard et al. [1] utilizaram o algoritmo APD para aproximar eficientemente a distância entre dois vértices de um grafo, que representa uma malha triangular, por sua distância geodésica. Tal veio na forma da proposição de um fator de alongamento (*stretch factor*), componente definido como a razão máxima entre o comprimento total de um caminho e o respectivo caminho mais curto, contabilizado pelo número de vértices (ou nós) que os compõem [12]. Valores pequenos para esse componente demonstram que a aproximação foi bem-feita.

A malha triangular referida em [1] representa uma discretização de uma superfície bi-dimensional mergulhada num espaço \mathbb{R}^d , com d constante, utilizando um conjunto de vértices V , um conjunto de arestas E , e um conjunto de faces triangulares F , que gera um grafo associado $G(V, E, F)$. Este grafo foi definido de forma que cada aresta é adjacente a um ou dois triângulos, e triângulos incidentes a um vértice qualquer podem ser ciclicamente ordenados em torno deste vértice.

Para minimizar os efeitos da discretização, os pesquisadores optaram por considerar que este grafo como sendo planar e conexo, ou seja, suas arestas são representadas em um plano sem cruzamento entre elas, e pode ser descrito um caminho (uma sequência de vértices) entre qualquer par de vértices do grafo [8]. Além disso, considerou-se que o grafo é ponderado, com valores positivos associados a suas arestas.

Para o caso apresentado, considerando um número inteiro k , o objetivo seria encontrar um conjunto $S = \{s_1, \dots, s_k\}$ de vértices de G que minimizam o fator de alongamento, definido por:

$$\max_{(p,q) \in V, p \neq q} \min_{s_i \in S} \frac{d(p, s_i) + d(s_i, q)}{d(p, q)}$$

onde $d()$ representa a caminho geodésico mais curto ente dois vértices p e q quaisquer.

4.1. Definição do problema

Ao estudar formas isométricas (como tecidos ou o corpo humano), descobre-se que as mesmas se deformam de forma quase isométrica, pois aplicando-se um alongamento suficientemente grande, as superfícies sofrem danos (como rasgos ou feridas). A análise de tais formas costuma ser feita através da comparação entre as geodésicas de vários pares de vértices para definir a quantidade de alongamento que possa ser aplicado à mesma.

A problemática trabalhada no grafo associado às malhas triangulares formadas nessas superfícies partiu inicialmente de dois conhecidos problemas na área da teoria dos grafos:

- Problema de caminho mínimo de origem única (*single-source shortest path* ou CMOU) – problema onde, dado um vértice de origem s num grafo ponderado G , deve-se encontrar o caminho mais curto entre este vértice e todos os demais vértices do grafo [13];
- Problema de caminho mínimo de todos os pares (*all-pair shortest path* ou CMTP) - problema onde deve-se calcular o caminho mais curto entre dois vértices, para todos os pares de vértices possíveis no grafo [14].

Usualmente, tais problemas podem ser resolvidos através da aplicação do algoritmo de Dijkstra [3], que resolve ambos com tempo $O(n \log n)$ e $O(n^2 \log n)$ respectivamente, com n sendo o número de vértices do grafo. Considerando a complexidade das malhas triangulares encontradas no mundo, o uso do CMTP acaba se tornando complicado, pois alguns algoritmos acabam por levar um tempo estrondosamente grande para resolvê-lo.

No intuito de diminuir a complexidade do problema final, foi considerado um problema onde uma estrutura de dados seria computada previamente para aproximar a distância entre dois vértices, sendo este chamado de Problema de caminho mínimo aproximado para qualquer par. A forma citada para a resolução deste problema é selecionar um conjunto de k raízes (vértices pertencentes ao

grafo), resolver o CMOU em cada uma, e com os resultados aproximar distâncias geodésicas parelhas.

Dado um par de vértices p e q , a distância mínima entre eles será dada pelo valor mínimo das distâncias entre estes e uma raiz s_i , considerando todas as k raízes, tornando-se necessário encontrar um posicionamento ideal de raízes de forma a minimizar o fator de alongamento. A este problema, Lazard et al. deram o nome de problema de dilatação de caminho de centro k , sendo este o problema estudado em [1].

4.2. O uso do algoritmo APD

Para a seleção das k raízes necessárias, os pesquisadores propuseram a utilização do algoritmo APD, onde iniciando de um vértice qualquer, seria acrescentado iterativamente ao conjunto de raízes S o vértice com maior distância geodésica de uma raiz próxima já selecionada, até que k raízes tenham sido escolhidas. Em um grafo G , esta distância é aproximada com sendo a distância mínima entre dois vértices através de uma raiz, considerando todas as k raízes.

Com isso, através do uso do algoritmo em k raízes, Lazard et al propuseram um fator de alongamento cujo resultado seria tão eficiente quanto o fator de alongamento mínimo obtido através de um conjunto ótimo de raízes. Tal valor, denominado por eles como \mathcal{F}_{FPS} , seria tal que:

$$\mathcal{F}_{FPS} \leq 2r_e^2(\mathcal{F}^* + 1) + 8r_e + 1$$

onde \mathcal{F}^* é o fator de alongamento obtido pelo conjunto ótimo de raízes, e r_e é a razão entre a aresta mais longa e a mais curta de G . O valor de r_e afeta diretamente os fatores de alongamento, sendo que se este for muito grande, os fatores também se tornam grandes.

Para comprovar este resultado, dois teoremas devem ser provados; um que prova o valor do fator de alongamento proposto, e outro que mostra que o Problema de dilatação de caminho de centro k é NP-completo. As definições e

provas dos teoremas são apresentadas a seguir, conforme originalmente mostradas no trabalho de Lazard et al.

4.3. Fator de alongamento

Inicialmente, considere um grafo conexo planar cujas arestas possuem comprimento definido num intervalo finito e positivo $[l_{min}l_{max}]$, e seja r_e a razão $\frac{l_{max}}{l_{min}}$. Dadas k raízes s_1, \dots, s_k no grafo, faça s_p denotando a raiz mais próxima ao vértice p , e $d(p, s_i, q)$ denotando o caminho de comprimento mais curto de p a q através de qualquer raiz s_1, \dots, s_k , representado por $\min_i(d(p, s_i) + d(s_i, q))$.

Seja s_1^*, \dots, s_k^* o conjunto de raízes que minimiza o fator de alongamento $F^* = \max_{p,q} \frac{d(p, s_i^*, q)}{d(p, q)}$ e, além disso, seja s'_1, \dots, s'_k um conjunto de raízes que minimiza $\max_p d(p, s_p)$. O conjunto de s_i^* é uma solução ótima para o problema de dilatação de caminho de centro k e o conjunto de s'_i é uma solução ótima para o problema de centro k (achar k raízes de forma que a distância máxima de qualquer ponto à raiz mais próxima é minimizado).

Dispostas estas considerações iniciais, enuncia-se o seguinte teorema:

Teorema 1: Seja s_1, \dots, s_k um conjunto de raízes retornado pelo algoritmo APD em um grafo conexo G cujas arestas possuem tamanho de no máximo r_e . Então:

$$\max_{p,q} \frac{d(p, s_i, q)}{d(p, q)} \leq 2r_e^2 \max_{p,q} \frac{d(p, s_i^*, q)}{d(p, q)} + 2r_e^2 + 8r_e + 1$$

Para a prova do teorema enunciado acima, será necessária a aplicação de alguns lemas adicionais, cuja combinação de resultados completará sua demonstração:

- Primeiro será mostrado que para qualquer conjunto de raízes, o fator de alongamento $\max_{p,q} \frac{d(p, s_i, q)}{d(p, q)}$ é determinado quando p e q forem adjacentes em um grafo G ;

- Com este resultado, o fator de alongamento será limitado em termos de $\max_p d(p, s_p)$;
- Em seguida, limita-se o tamanho do caminho entre dois vértices quaisquer através de uma determinada raiz em termos do menor caminho que passe entre estes vértices através de qualquer raiz.
- Por fim, o fator de alongamento será limitado para qualquer conjunto de raízes em termos do fator de um conjunto ótimo de raízes para o problema de centro k .

Lema 1: Para quaisquer raízes s_1, \dots, s_k e qualquer vértice q dado em G , a razão máxima $\max_p \frac{d(p, s_i, q)}{d(p, q)}$ é encontrada para algum p adjacente a q em G . Disso, segue que a razão máxima $\max_{p, q} \frac{d(p, s_i, q)}{d(p, q)}$ é encontrada para alguns p e q adjacentes em G .

Prova: Para demonstrar por contradição, faça q ser qualquer vértice fixo e p um vértice não-adjacente que determine $\max_p \frac{d(p, s_i, q)}{d(p, q)}$, e tais que dentre todos os vértices que determinem esse valor, o caminho mais curto de p a q possui o menor número de arestas.

Seja \tilde{p} o vizinho imediato de p no caminho mais curto de p a q . Como dito, $d(\tilde{p}, s_j, q)$ denota o caminho de tamanho mais curto entre \tilde{p} e q através de quaisquer raízes s_1, \dots, s_k . Tendo $d(\tilde{p}, s_j, q) \geq d(p, s_i, q) - l$, e dividindo por $d(\tilde{p}, q) = d(p, q) - l$, obtém-se:

$$\frac{d(\tilde{p}, s_j, q)}{d(\tilde{p}, q)} \geq \frac{d(p, s_i, q)}{d(p, q) - l} - \frac{l}{d(\tilde{p}, q)}$$

Por outro lado, multiplicando $d(p, q) = d(\tilde{p}, q) + l$ por $\frac{d(p, s_i, q)}{d(p, q)d(\tilde{p}, q)}$, obtém-se:

$$\frac{d(p, s_i, q)}{d(p, q) - l} = \frac{d(p, s_i, q)}{d(p, q)} + \frac{ld(p, s_i, q)}{d(\tilde{p}, q)d(p, q)}$$

E daí:

$$\frac{d(\tilde{p}, s_j, q)}{d(\tilde{p}, q)} \geq \frac{d(p, s_i, q)}{d(p, q)} + \frac{l}{d(\tilde{p}, q)} \left(\frac{d(p, s_i, q)}{d(p, q)} - 1 \right) \geq \frac{d(p, s_i, q)}{d(p, q)}$$

Isto contradiz a hipótese inicial, visto que ou a inequação acima é estrita e $\frac{d(p, s_i, q)}{d(p, q)}$ não é máximo, ou a igualdade vale e o caminho mais curto entre \tilde{p} e q tem menos arestas que o caminho mais curto entre p e q .

Lema 2: Para quaisquer raízes s_1, \dots, s_k , tem-se:

$$\frac{2}{l_{\max}} \max_p d(p, s_p) - 1 \leq \max_{p, q} \frac{d(p, s_i, q)}{d(p, q)} \leq \frac{2}{l_{\min}} \max_p d(p, s_p) + 1$$

Prova: Para o limite superior tem-se $d(p, s_i, q) \leq d(p, s_p) + d(s_p, q) \leq 2d(p, s_p) + d(p, q)$. Por isso tem-se, $\frac{d(p, s_i, q)}{d(p, q)} \leq \frac{2}{d(p, q)} d(p, s_p) + 1$, que é assegurado para quaisquer vértices p e q , e portanto para aqueles que determinam o valor máximo para $\frac{d(p, s_i, q)}{d(p, q)}$. Além disso, $d(p, q) \geq l_{\min}$ e $d(p, s_p) \leq \max_p d(p, s_p)$. Consequentemente:

$$\max_{p, q} \frac{d(p, s_i, q)}{d(p, q)} \leq \frac{2}{l_{\min}} \max_p d(p, s_p) + 1$$

Para o limite inferior, sabe-se que pela desigualdade triangular, para qualquer i , $d(q, s_i) \geq d(p, s_i) - d(p, q)$, e adicionando $d(p, s_i)$ em ambos os membros, obtêm-se $d(p, s_i) + d(q, s_i) \geq 2d(p, s_i) - d(p, q)$. Pela definição de s_p , $d(p, s_i) \geq d(p, s_p)$ vale para qualquer i , daí $d(p, s_i) + d(q, s_i) \geq 2d(p, s_p) - d(p, q)$, e para o i tal que $d(p, s_i) + d(q, s_i)$ é mínimo, tem-se $d(p, s_i, q) \geq 2d(p, s_p) - d(p, q)$.

Dividindo por $d(p, q)$, obtem-se $\frac{d(p, s_i, q)}{d(p, q)} \geq \frac{2}{d(p, q)} d(p, s_p) - 1$, que vale para qualquer p e q e em especial para o p que maximiza $d(p, s_p)$ (seja este vértice denotado por \tilde{p}). Desta forma, tem-se que $\frac{d(\tilde{p}, s_i, q)}{d(\tilde{p}, q)} \geq \frac{2}{d(\tilde{p}, q)} \max_p d(p, s_p) - 1$, que

vale para qualquer q , e em particular para aquele que determina $\max_q \frac{d(\tilde{p}, s_i, q)}{d(p, q)}$. Pelo Lema 1, o máximo é encontrado para um q que seja adjacente a \tilde{p} em G , então, para este q , tem-se $\frac{2}{d(\tilde{p}, q)} \geq \frac{2}{l_{max}}$. E disso, tem-se:

$$\max_{p, q} \frac{d(p, s_i, q)}{d(p, q)} \geq \max_{p, q} \frac{d(\tilde{p}, s_i, q)}{d(\tilde{p}, q)} \geq \frac{2}{l_{max}} \max_p d(p, s_p) - 1$$

Lema 3: Para quaisquer raízes s_1, \dots, s_k , e vértices u e v , tem-se:

$$d(u, s_u) + d(s_u, v) \leq d(u, s_i, v) + 2d(u, v)$$

Prova: Seja s_i a raiz que determina $d(u, s_i, v) = \min_o (d(u, s_i) + d(s_i, v))$. Por definição $d(u, s_u) \leq d(u, s_i)$, portanto só é necessário mostrar que $d(v, s_u) \leq d(s_i, v) + 2d(u, v)$. Usando a desigualdade triangular duas vezes, tem-se:

$$d(v, s_u) \leq d(v, u) + d(u, s_u) \leq d(v, u) + d(u, s_i) \leq d(v, u) + d(u, v) + d(u, s_i)$$

Lema 4: Sejam s_1, \dots, s_k um conjunto de raízes retornados pelo algoritmo APD, e s'_1, \dots, s'_k um conjunto ótimo de raízes para o problema de centro k . Então:

$$\max_{p, q} \frac{d(p, s_i, q)}{d(p, q)} \leq 2r_e \max_{u, v} \frac{d(u, s'_i, v)}{d(u, v)} + 6r_e + 1$$

Prova: Como s_1, \dots, s_k é um conjunto de raízes retornado pelo algoritmo APD, esta escolha de raízes provém uma 2-aproximação para o problema de centro k em comparação com uma solução ótima s'_1, \dots, s'_k , ou seja, $\max_p d(p, s_p) \leq 2 \max_p d(p, s'_p)$.

Por definição, $d(p, s_i, q)$ é o mínimo sobre todas as raízes fixas s_i de $d(p, s_i) + d(s_i, q)$, então $d(p, s_i, q) \leq d(p, s_p) + d(s_p, q)$, e ainda mais, pela desigualdade triangular, $d(s_p, q) \leq d(s_p, p) + d(p, q)$, sendo $d(p, s_i, q) \leq 2d(p, s_p) + d(p, q)$. Entretanto, $d(p, s_p) \leq \max_u d(u, s_u)$, que é menor ou igual a $2 \max_u d(u, s'_u)$ pela propriedade da 2-aproximação. Para um melhor

entendimento, denota-se o vértice que determina $2 \max_u d(u, s'_u)$ por u , tendo-se então $d(p, s_i, q) \leq 4d(u, s'_u) + d(p, q)$.

Pela desigualdade triangular, $d(u, s'_u) \leq d(u, v) + d(v, s'_u)$ para qualquer v , e assim, $2d(u, s'_u) \leq d(u, v) + d(v, s'_u) + d(u, s'_u)$. Pelo Lema 3, $2d(u, s'_u) \leq 3d(u, v) + d(u, s'_i, v)$, e $d(p, s_i, q) \leq 2d(u, s'_i, v) + 6d(u, v) + d(p, q)$. Dividindo esta expressão por $d(p, q)$ e multiplicando por $\frac{d(u, v)}{d(u, v)}$, obtém-se:

$$\frac{d(p, s_i, q)}{d(p, q)} \leq 2 \frac{d(u, v)}{d(p, q)} \frac{d(u, s'_i, v)}{d(u, v)} + 6 \frac{d(u, v)}{d(p, q)} + 1$$

onde esta inequação é válida para qualquer par distinto de p e q , e qualquer v distinto de u fixo. Sendo assim, ela vale para p e q particulares que determinam $\max_{p, q} \frac{d(p, s_i, q)}{d(p, q)}$, e v que determina $\max_v \frac{d(u, s'_i, v)}{d(u, v)}$. Pelo Lema 2, este v é vizinho de u e satisfaz $d(u, v) \leq l_{max}$, e, sendo $d(p, q) \geq l_{min}$ para qualquer par distinto de p e q , e $\max_v \frac{d(u, s'_i, v)}{d(u, v)} \leq \max_{u, v} \frac{d(u, s'_i, v)}{d(u, v)}$, tem-se:

$$\max_{p, q} \frac{d(p, s_i, q)}{d(p, q)} \leq 2 \frac{l_{max}}{l_{min}} \max_{u, v} \frac{d(u, s'_i, v)}{d(u, v)} + 6 \frac{l_{max}}{l_{min}} + 1$$

Com estes resultados garantidos, prova-se então o Teorema 1:

Prova (Teorema 1): Utilizando as notações pré-definidas, pelo Lema 4 tem-se:

$$\max_{p, q} \frac{d(p, s_i, q)}{d(p, q)} \leq 2r_e \max_{u, v} \frac{d(u, s'_i, v)}{d(u, v)} + 6r_e + 1$$

Utilizando o limite superior mostrado no Lema 2 em $\max_{u, v} \frac{d(u, s'_i, v)}{d(u, v)}$, obtém-se:

$$\max_{p, q} \frac{d(p, s_i, q)}{d(p, q)} \leq 2r_e \left(\frac{2}{l_{min}} \max_p d(p, s'_p) + 1 \right) + 6r_e + 1$$

Por definição, s'_1, \dots, s'_k é um conjunto ótimo de raízes para o problema de centro k , ou seja, $argmin_{s_1, \dots, s_k} \max_p d(p, s_p)$. Então, $\min_{s_1, \dots, s_k} \max_p d(p, s_p) =$

$\max_p d(p, s'_p) \leq \max_p d(p, s_p^*)$, e aplicando o limite inferior dado pelo Lema 2 para s_1^*, \dots, s_k^* obtém-se:

$$\frac{2}{l_{max}} \max_p d(p, s_p^*) - 1 \leq \max_{p,q} \frac{d(p, s_i^*, q)}{d(p, q)}$$

E por fim:

$$\begin{aligned} \max_{p,q} \frac{d(p, s_i, q)}{d(p, q)} &\leq 2r_e \left(\frac{l_{max}}{l_{min}} \left(\max_{p,q} \frac{d(p, s_i^*, q)}{d(p, q)} + 1 \right) + 1 \right) + 6r_e + 1 \\ &\leq 2r_e^2 \max_{p,q} \frac{d(p, s_i^*, q)}{d(p, q)} + 2r_e^2 + 8r_e + 1 \end{aligned}$$

4.4. Complexidade do Método

Comprovado o valor do fator de alongamento, agora é necessário comprovar a complexidade do método. Para o modelo em questão, considerou-se a complexidade de um problema de dilatação de caminho de centro k aplicado em grafos triangulares, ou seja, o cálculo de um conjunto ótimo de raízes que minimiza o fator de alongamento.

Deve-se notar que o problema será NP se, para qualquer conjunto de k raízes, o fator de alongamento é computado em tempo polinomial. A dificuldade do problema será mostrada ao reduzir o problema decisivo relacionado a encontrar uma cobertura de vértices com cardinalidade mínima em um grafo planar de grau máximo igual a três.

O teorema enunciado a seguir mostra que a versão decisiva deste problema é NP-Completa para grafos triangulares:

Teorema 2: Dado um grafo $G = (V, E, F)$, um número inteiro k , e um valor real ξ , é NP-Completo determinar se existe um conjunto $S = \{s_1, \dots, s_k\}$ de k raízes tais que o valor do fator de alongamento $\max_{p,q} \min_i \frac{d(p, s_i) + d(s_i, q)}{d(p, q)}$ é no máximo ξ .

Dois lemas devem ser considerados para realizar a redução citada anteriormente, que proverá a demonstração do teorema:

- O primeiro mostra que grafos planares podem ser imersos em grades de números inteiros (conforme apresentado em [5]).
- O segundo mostra que o grafo só possui uma cobertura de vértices (ou seja, um conjunto de vértices tal que cada aresta do grafo incide em pelo menos um vértice do conjunto) de tamanho k , se sua imersão planar tem uma cobertura de tamanho $k + m$, com $m = \sum_{e \in E} k_e$.

Lema 5: Um grafo planar $G = (V, E)$ com grau máximo igual a 4 pode ser imerso em um plano usando uma área $O(|V|^2)$ de tal forma que seus vértices possuem coordenadas inteiras, e suas arestas são desenhadas como segmentos de linha poligonais que se encontram sobre uma grade de inteiros (ou seja, cada aresta consiste de um ou mais segmentos de linha que se encontram sobre linhas da forma $x = i$ ou $y = j$, onde i e j são inteiros).

Considere um grafo planar G com grau máximo igual a 3, e seja G_r uma imersão planar de G (isto é, uma transcrição do grafo em um plano) de acordo com o Lema 5, sendo acrescentado para cada aresta $e \in E$, um número par de $2k_e$ nós auxiliares com coordenadas meio-inteiras e tais que cada aresta resultante em G_r possui comprimento igual a 1 ou $\frac{1}{2}$. A grade de meio-inteiros é considerada para assegurar que um número par de nós auxiliares seja adicionado em cada aresta de G , fazendo com que as arestas resultantes de G_r tenham comprimento de no máximo 1.

Para uma aresta $uv \in E$, use $cam(u, v)$ para denotar o caminho em G_r que substitui a aresta uv , e chame os pontos terminais deste caminho de nós regulares.

Lema 6: G possui uma cobertura de vértices de tamanho k se, e somente se, G_r possui uma cobertura de vértices de tamanho $k + m$.

Prova: Qualquer cobertura de vértices C de G com tamanho igual a k pode ser estendida para uma cobertura de vértices de tamanho $k + m$ em G_r ao incluir todo nó auxiliar em $cam(u, v)$ para cada aresta $uv \in E$.

Seja C_r uma cobertura para G de tamanho $m + k$ e suponha que existe um caminho $cam(u, v)$ onde nem u e nem v pertencem a C_r . Isto significa que ao menos $k_{uv} + 1$ nós auxiliares de $cam(u, v)$ pertencem a C_r , para que todas as arestas deste caminho sejam cobertas. Porém, ao usar apenas os k_{uv} nós auxiliares de $cam(u, v)$ e adicionar u e v a C_r , a cobertura de vértices continua do mesmo tamanho mas contendo os pontos terminais do caminho citado.

Partindo disso, é possível construir uma cobertura de vértices C_r de tamanho $k + m$ para G_r que inclua ao menos um dos pontos terminais de cada caminho $cam(u, v)$, $uv \in E$, onde tal cobertura C_r será uma cobertura para G restrita aos seus nós regulares. Como é necessário um número mínimo de k_{uv} nós auxiliares para cobrir qualquer caminho $cam(u, v)$, $uv \in E$, mesmo que ambos u e v pertençam a cobertura, o número de nós regulares selecionados é de no máximo k .

Considerando o resultado anterior, substitui-se cada aresta de G_r por um dispositivo q como mostrado na figura a seguir, e denota-se o grafo resultante por G' . Assim, prova-se o Teorema 2:

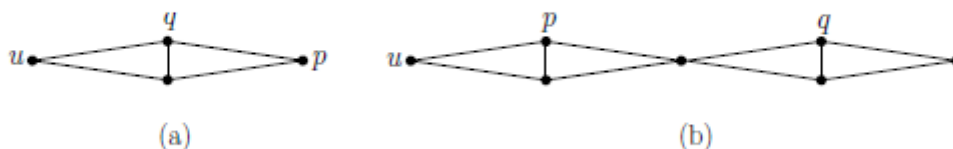


Figura 3: Dispositivos gráficos utilizados na prova do Teorema 2

Prova (Teorema 2): Considere o grafo G' construído a partir de um grafo planar G com grau máximo 3 e dispositivos tais que $\frac{|au|+|bu|}{|ab|} = \xi \geq 3$. Este grafo pode ser visto como uma união de triângulos e, portanto, como uma malha triangular. Deve ser mostrado que G possui uma cobertura de tamanho k se, e somente se, G' tem $k + m$ raízes tais que o fator de alongamento é no máximo ξ , o que faz com que o problema da cobertura de vértices possa ser reduzido ao problema inicialmente enunciado em tempo polinomial.

Pelo Lema 6, se G possui uma cobertura de tamanho k , então G_r possui uma cobertura de tamanho $k + m$, lembrando que esta cobertura de G_r pode ser obtida adicionando todos os nós auxiliares de cada aresta de G a sua cobertura de vértices. Escolhendo esta cobertura de $k + m$ nós como raízes de G' , e considerando um par de nós p e q , três casos podem ser analisados:

- p e q pertencem ao mesmo dispositivo

Sejam u, v, a e b os nós desse dispositivo e suponha que sem perda de generalidade, u é selecionado como raiz. Então, $\frac{d(p,u)+d(u,q)}{d(p,q)}$ é igual a 1 se p ou q coincidirem com u (e por definição igual a ξ se $(p, q) = (a, b)$), e igual a 3 se $p = v$ e $q = a$ ou $q = b$. Como por definição do dispositivo $\xi \geq 3$, o valor máximo de $\min_i \frac{d(p, s_i^*) + d(s_i^*, q)}{d(p, q)}$ é ξ para todos os pares (p, q) do dispositivo.

- p e q pertencem a dispositivos ϱ_1 e ϱ_2 adjacentes

Sejam $\{u_1, a_1, b_1, v_1, a_2, b_2, v_2\}$ os nós dos dois dispositivos. Caso o nó v_1 (que deve pertencer a ambos), p ou q forem selecionados como raiz, o caso recai para o Caso 1, por simetria, e fornece o mesmo limite para ξ . Por outro lado, sendo u_1 e v_2 raízes, para quaisquer dois nós de dois dispositivos diferentes a razão máxima será 2, o que mostra que ξ ainda é o valor máximo.

- p e q não pertencem ao mesmo dispositivo, nem a dois dispositivos adjacentes

Neste caso pelo menos um dos nós no caminho mais curto entre p e q é selecionado como raiz, e então seu caminho mais curto aproximado será igual ao caminho geodésico mais curto e $\min_i \frac{d(p, s_i^*) + d(s_i^*, q)}{d(p, q)} = 1$.

Assim, prova-se que o fator de alongamento de G' é ξ para as $k + m$ raízes selecionadas. Agora, mostra-se que se G' possui $k + m$ raízes de forma que o fator de alongamento é no máximo ξ , então G tem uma cobertura de vértices de tamanho k .

Cada dispositivo deve conter pelo menos uma raiz, visto que os vértices a e b de um dispositivo sem raiz provém $\min_i \frac{d(a, s_i^*) + d(s_i^*, b)}{d(a, b)} > \xi$. Então, para cada dispositivo em G' , se um dos vértices u ou v for raiz, seleciona-se o vértice correspondente em G_r e se um dos vértices a ou b for raiz, seleciona-se o ponto terminal da aresta de G_r correspondente ao dispositivo. Assim, pelo menos um vértice do grafo G_r é selecionado, e então G_r tem uma cobertura de vértice de tamanho $k + m$, o que pelo Lema 6 implica em que G possua uma cobertura de tamanho k .

5. Conclusões

O trabalho proposto por Lazard et al demonstrou uma interessante aplicação do algoritmo APD, utilizando-o com conceitos de grafos para obter um fator de alongamento eficiente para o cálculo de distâncias geodésicas. Entretanto, como citado pelos próprios autores, ainda não há provas teóricas de que o fator apresentado possua eficiência superior a um fator obtido por um conjunto ótimo de raízes.

Dessa forma, para o futuro, espera-se a comprovação da eficiência de tal elemento em aplicações que envolvam o cálculo de geodésicas. Por se tratar de uma proposição extremamente recente, é esperado que haja alguma demora para que o referido resultado seja publicado, porém, a utilização do mesmo em outros projetos pode ser realizada com a finalidade de testá-lo, a fim de comprovar sua eficácia.

Referências Bibliográficas

- [1] **Lazard, S., Kamousi, P., Maheshwari, A., Wuhrer, S.** Analysis of Farthest Point Sampling for Approximating Geodesics in a Graph. *Computational Geometry*, ed. 57, p: 1-13, Elsevier, 2016.
- [2] **Eldar, Y., Lindebaum, M., Porat, M., Zeevi, Y. Y.** The Farthest Point Strategy for Progressive Image Sampling. *Transactions on Image Processing*, vol. 6 n° 9 p: 1305-1315, 1997.
- [3] **Dijkstra, E. W.** A note on two problems in connexion with graphs. *NumerischeMathematik*, vol. 1, p: 269-271, 1959
- [4] **Qin, W., Hu, Y., Sun, Y., & Yin, B.** An automatic multi-sample 3d face registration method based on thin plate spline and deformable model. In: *Multimedia and Expo Workshops, 2012 IEEE International Conference*, p: 453-458. IEEE, 2012.
- [5] **Valiant, L. G.** Universality Considerations on VSLI Circuits. *IEEE Transactions on Computers*, vol. 30, n° 2, p: 135-140, IEEE, 1981.
- [6] **Liu R., Jain V., Zhang H.** Sub-sampling for Efficient Spectral Mesh Processing. In: Nishita T., Peng Q., Seidel HP. (eds) *Advances in Computer Graphics. Lecture Notes in Computer Science*, vol 4035. Springer, Berlin, Heidelberg, 2006.
- [7] **Moening, C., Dodgson, N. A.** Intrinsic point cloud simplification. *Proc. 14th GrahiCon*, v. 14, p. 23, 2004.
- [8] **Feofiloff, P., Kohayakawa, Y., & Wakabayashi, Y.** Uma introdução sucinta à teoria dos grafos, 2011. Disponível em <http://www.ime.usp.br/~pf/teoriadosgrafos>. Acessado em: 17 de abril de 2017.
- [9] **Pokorny, P.** Geodesics Revisited. *Chaotic Modeling and Simulation*, 2012.
- [10] **Moening, C., Dodgson, N. A.** Fast marching farthest point sampling for implicit surfaces and point clouds. *Computer Laboratory Technical Report*, vol. 565, 2003.

- [11] **Moening, C., Dodgson, N. A.** Fast marching farthest point sampling (No. UCAM-CL-TR-562). University of Cambridge, Computer Laboratory, 2003.
- [12] **Busch, C., Magdon-Ismail, M., Xi, J.** Oblivious routing on geometric networks. In: *Proceedings of the seventeenth annual ACM symposium on Parallelism in algorithms and architectures*, p. 316-324, ACM, 2005.
- [13] **Thorup, M.** Undirected single-source shortest paths with positive integer weights in linear time. *Journal of the ACM (JACM)*, vol.46, n° 3, p. 362-394, 1999.
- [14] **Katz, G. J., Kider Jr, J. T.** All-pairs shortest-paths for large graphs on the GPU. In *Proceedings of the 23rd ACM SIGGRAPH/EUROGRAPHICS symposium on Graphics hardware*, Eurographics Association, p. 47-55, 2008.
- [15] **Bronstein, A. M., Bronstein, M. M., Kimmel, R.** Generalized multidimensional scaling: a framework for isometry-invariant partial surface matching. *Proceedings of the National Academy of Sciences*, vol. 103, n° 5, p. 1168-1172, 2006.
- [16] **Ruggeri, M. R., Saupe, D.** Isometry-invariant Matching of Point Set Surfaces. In: *3DOR*, p. 17-24, 2008.
- [17] **Bougleux, S., Peyré, G., Cohen, L. D.** Image compression with anisotropic triangulations. In: *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, p. 2343-2348, 2009.