

UNIVERSIDADE DE SÃO PAULO
Instituto de Ciências Matemáticas e de Computação

**Reconstrução de curvas por meio de características
robustas extraídas de imagens**

Projeto de Pesquisa apresentado à FAPESP
para solicitação de Bolsa de Pesquisa de Ini-
ciação Científica (IC).

Graduando: André Luís Mendes Fakhoury

Orientador: João do Espírito Santo Batista Neto

Brasil

2020

Resumo

Um dos desafios atuais da área de processamento de imagens e visão computacional é o mapeamento de características robustas entre espaços bidimensionais e tridimensionais. A simplificação deste processo pode ser feita através da redução da quantidade de informações, como, por exemplo, representar curvas a partir de pontos (denominados características robustas ou pontos importantes), de forma que estas possam ser reconstruídas através destes com um erro mínimo. Assim, contornos de imagens podem ser extraídos e analisados como curvas discretas e, com isso, reduzidos a finitos pontos importantes. Portanto, a pesquisa analisará a extração de pontos importantes de curvas discretas e a seguinte reconstrução de curvas a partir destes, por meio de curvas poligonais. Este estudo insere-se em umas das linhas de atuação do projeto temático *FAPESP 2019/07316-0*, que visa a reconstrução de faces humanas a partir de informações reduzidas do domínio.

Palavras-chave: imagens. curvas. características robustas. restauração.

Sumário

1	Introdução e justificativa	4
2	Objetivos	6
3	Conceitos básicos e fundamentos teóricos	6
3.1	Pré-processamento em imagens	6
3.1.1	Suavização	6
3.1.2	Binarização	7
3.1.3	Erosão e dilatação	8
3.1.4	Extração de contorno	9
3.2	Cálculo de curvatura	9
3.3	Características robustas	10
3.4	Reconstrução de curvas	10
3.4.1	O Operador Laplaciano e a representação diferencial	11
3.4.2	Reconstrução de curvas a partir das δ -coordenadas	13
3.5	Métricas para avaliação de similaridades entre contornos	13
3.5.1	Covering	14
3.5.2	Medida Baseada em contornos	15
3.5.3	Probabilistic Rand Index (PRI)	15
4	Materiais e métodos	15
5	Plano de trabalho e cronograma	17
	Referências	19

1 Introdução e justificativa

Uma tarefa ainda desafiadora em processamento de imagens e visão computacional é a reconstrução de elementos de uma cena a partir de pontos de controle.

Do ponto de vista geométrico, esta tarefa pode ser vista como a identificação de características robustas e posterior aplicação de alguma técnica que permita reconstruir curvas a partir de tais características (SORKINE, 2006).

Características robustas, termo cunhado por Ian Porteus, são características de uma superfície que se mantém quando esta se deforma (IZUMIYA et al., 2015). Como exemplo de características robustas em superfície, têm-se curvas parabólicas (formadas pelo contato da superfície com linhas e planos). Em curvas bidimensionais, suas características robustas (ou pontos importantes) são os pontos críticos da curvatura (máximos e mínimos), pontos de inflexão, ou que intersectam o eixo-x.

Em imagens, características robustas podem ser extraídas empregando-se um conjunto de técnicas costumeiramente utilizadas em visão computacional e processamento de imagens.

De acordo com Gonzalez e Woods (2018), imagens, a partir de uma visão matemática, podem ser vistas como funções do tipo $f(x, y)$, em que x e y designam a localização dos pixels. Em imagens monocromáticas, por exemplo, os valores de f são números inteiros (que variam de 0 a 255) que representam a quantidade de luz em determinado pixel (x, y) . Já em imagens do tipo *RGB*, os valores de $f(x, y)$ são uma trinca, representando, respectivamente, a quantidade de vermelho, verde e azul de cada pixel.

De uma imagem $f(x, y)$ é possível extrair contornos a partir dos quais podem ser geradas as características robustas. Em processamento de imagens, contornos são curvas que permitem delimitar um objeto (vide figura 1). Este processo, invariavelmente, requer a simplificação da imagem, de forma a eliminar informações desnecessárias. Este pré-processamento consiste de técnicas como binarização, segmentação e remoção de ruído.

Curvas na dimensão \mathbb{R}^2 são definidas por funções $F : A \rightarrow \mathbb{R}^2$, com F de classe C^2 e $A \subseteq \mathbb{R}$ intervalo fechado - ou seja, funções de uma variável real a valores em \mathbb{R}^2 . Nestas, cada valor $t \in A$ é associado a um vetor bidimensional. Em imagens, podem ser visualizadas curvas em contornos de figuras.

A partir de uma amostragem finita de pontos pode-se reconstruir curvas e superfícies. Existem diversos algoritmos com tal finalidade, podendo ser baseados majoritariamente em técnicas interpolatórias, funções implícitas ou geração de malhas (SORKINE, 2006).

Os baseados em técnicas interpolatórias normalmente se baseiam na triangulação

Figura 1 – Formas de figuras em que os contornos são facilmente extraídos



Fonte: Costa e Jr (2009).

de Delaunay (BERG et al., 2000) e respectiva interpolação dos pontos e, por isso, tem custo computacional alto. Já as técnicas baseadas em funções implícitas realizam a representação das curvas e superfícies por um conjunto de funções implícitas e, portanto, permitem a realização de operações padrões matemáticas com estas funções, porém impossibilitam o percurso sobre a superfície - porém, são mais adequadas que técnicas interpolatórias se a nuvem de pontos tiver ruídos. Os métodos baseados em malhas realizam a reconstrução de superfícies a partir de uma quantidade definida de malhas (quanto mais malhas, mais próxima a superfície reconstruída fica da original). Em casos bidimensionais, não são consideradas malhas, mas sim curvas poligonais.

A extração de características robustas a partir de imagens, proposta neste projeto de Iniciação Científica, é parte de uma das linhas do projeto temático (FAPESP 2019/07316 – 0), intitulado *Teoria de singularidades e aplicações a geometria diferencial, equações diferenciais e visão computacional*. Das 4 linhas de atuação do projeto, uma diz respeito à reconstrução de faces humanas a partir de informações reduzidas do domínio. Por exemplo, deseja-se reconstruir uma face humana (nariz, boca, olhos dentre outros elementos) a partir de uma versão simplificada desta (por exemplo, uma caricatura representada por curvas simplificadas da face). O ponto de partida para esta reconstrução é a identificação das características robustas.

2 Objetivos

O objetivo principal deste projeto de pesquisa é extrair características robustas em \mathbb{R}^2 para reconstrução de curvas com alta precisão a partir de imagens.

A realização desta tarefa implica na observância dos seguintes objetivos específicos:

- Pré-processar as imagens visando a eliminação de ruídos, binarização e consequente segmentação;
- Extrair atributos de formas das imagens (contorno e cálculo da curvatura);
- Determinar as características robustas por análise da curvatura;
- Reconstruir as formas a partir das características robustas, por meio de curvas poligonais e operadores Laplacianos como sugerido por Sorkine (2006) e
- Aferir a qualidade da reconstrução a partir da curva original.

3 Conceitos básicos e fundamentos teóricos

A execução deste projeto requer o conhecimento em áreas de conhecimentos distintos, que vão do processamento de imagens à geometria diferencial. Nesta seção, serão detalhados os principais conceitos necessários ao desenvolvimento do trabalho.

3.1 Pré-processamento em imagens

Para se extrair contornos em imagens reais, muitas vezes é necessário realizar operações de pré-processamento. Tais processos são de naturezas diversas, envolvendo desde a conversão de modelos de cores à segmentação. As principais atividades são descritas a seguir.

3.1.1 Suavização

A partir de imagem em escalas de cinza, realiza-se a suavização da imagem, com o objetivo de eliminar ruídos. O método mais simples de suavização é a aplicação de filtro gaussiano (SZELISKI, 2011).

O filtro gaussiano pode ser realizado a partir da seguinte equação:

$$G(x) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2}{2\sigma^2}},$$

em que σ representa o desvio padrão, com média 0 (ou seja, é uma distribuição gaussiana centrada na origem). Esta distribuição é então utilizada como uma função para propagar os pontos da imagem a partir de uma convolução. Quanto maior o valor de σ , maior o efeito de borrimento produzido na imagem.

Porém, o filtro gaussiano possui a propriedade de borrar a imagem de forma indistinta, tanto em regiões de baixa como de alta frequência. Isso pode provocar, por exemplo, o deslocamento de bordas que definem contornos importantes dos objetos. Neste projeto, é importante extrair contornos precisos, que sejam fiéis ao real contorno dos objetos presentes na imagem.

Para situações que requerem a preservação de bordas e ao mesmo tempo suavização em regiões de baixa frequência, é possível aplicar o filtro de difusão anisotrópica Perona Malik (BORROTO-FERNÁNDEZ; MECIAS; HIDALGO, 2013). Seja $\Omega \subset \mathbb{R}^2$ e $I(x, y, t)$ em Ω a imagem suavizada no tempo t (sendo a imagem original quando $t = 0$). A difusão anisotrópica é definida como:

$$\frac{\partial I}{\partial t} = \text{div}(c(x, y, t)\nabla(I(x, y, t))) = \nabla(c(x, y, t)) \cdot \nabla(I(x, y, t)) + c(x, y, t)\nabla(I(x, y, t)),$$

sendo ∇ o gradiente, $\text{div}(\dots)$ o operador divergência e escolhendo o coeficiente de difusão $c(x, y, t) = g(|\nabla I(x, y)|)$. Com isso, Perona e Malik (1990) propuseram duas funções para a função de difusão $g(|\nabla I(x, y)|)$:

$$g_1(|\nabla I(x, y)|) = \exp\left\{-\left[\frac{|\nabla I(x, y)|}{k}\right]^2\right\}$$

$$g_2(|\nabla I(x, y)|) = \frac{1}{1 + \left(\frac{|\nabla I(x, y)|}{k}\right)^2}$$

Tais funções, em regiões de alta frequência, identificadas pela computação do gradiente, fazem com que o filtro reduza o seu efeito de filtragem. Na prática, produz uma imagem em que as bordas são preservadas em suas posições originais, ao mesmo tempo que suaviza áreas de baixa frequência.

3.1.2 Binarização

Uma maneira eficiente para se obter o contorno de um objeto é transformar a imagem em uma apresentação binária. Uma imagem binária possui pixels de valores 0 ou 1, em que 1 representa posições do plano superior e 0 representa o plano de fundo (BURGER; BURGE, 2016).

A binarização pressupõe a existência de um parâmetro conhecido como limiar. Todo pixel com valor maior ou igual terá valor binário 1, enquanto os de valor menor terão valor binário 0.

A questão mais importante deste processo é a escolha automática do limiar. Para isso, pode-se utilizar o método de Otsu (OTSU, 1979). O algoritmo de Otsu calcula o valor de um limiar t que minimiza a variância ponderada intra-classes, dada pela relação:

$$\sigma_w^2(t) = q_1(t)\sigma_1^2(t) + q_2(t)\sigma_2^2(t)$$

sendo as probabilidades de cada classe (objeto e fundo, por exemplo) estimadas por:

$$q_1(t) = \sum_{i=1}^t P(i) \quad \& \quad q_2(t) = \sum_{i=t+1}^I P(i)$$

e as médias de ambas as classes computadas como:

$$\mu_1(t) = \sum_{i=1}^t \frac{iP(i)}{q_1(t)} \quad \& \quad \mu_2(t) = \sum_{i=t+1}^I \frac{iP(i)}{q_2(t)}$$

Finalmente, as variâncias para cada classe podem ser assim calculadas:

$$\sigma_1^2(t) = \sum_{i=1}^t [i - \mu_1(t)]^2 \frac{P(i)}{q_1(t)} \quad \& \quad \sigma_2^2(t) = \sum_{i=t+1}^I [i - \mu_2(t)]^2 \frac{P(i)}{q_2(t)}$$

Para uma imagem em nível de cinza, basta percorrer a faixa de valores $t = 0, \dots, 255$ e escolher aquele que minimiza $\sigma_w^2(t)$.

3.1.3 Erosão e dilatação

Outro ponto importante da pré-segmentação diz respeito a operadores elementares da morfologia matemática (SERRA, 1984) - a erosão e a dilatação.

De acordo com Regazzoni, Foresti e Venetsanopoulos (1994), a dilatação é feita com o objetivo de preencher pequenos buracos e ampliar a figura original, de modo a eliminar pequenas imperfeições formadas por outros pré-processamentos aplicados ou pela falta de informações na imagem original. Formalmente, sendo A um conjunto de pixels, B um elemento estruturante e z os valores de primeiro plano, a operação de dilatação $A \oplus B$ é definida como:

$$A \oplus B = \{z | (\hat{B})_z \cap A \neq \emptyset\}$$

Já a erosão é feita com o objetivo de “diminuir” o objeto, podendo expandir “buracos”, remover pequenas áreas e separar regiões próximas. Formalmente, a operação de erosão $A \ominus B$ entre o conjunto A e o elemento estruturante B é definida como:

$$A \ominus B = \{z | (B)_z \subseteq A\}$$

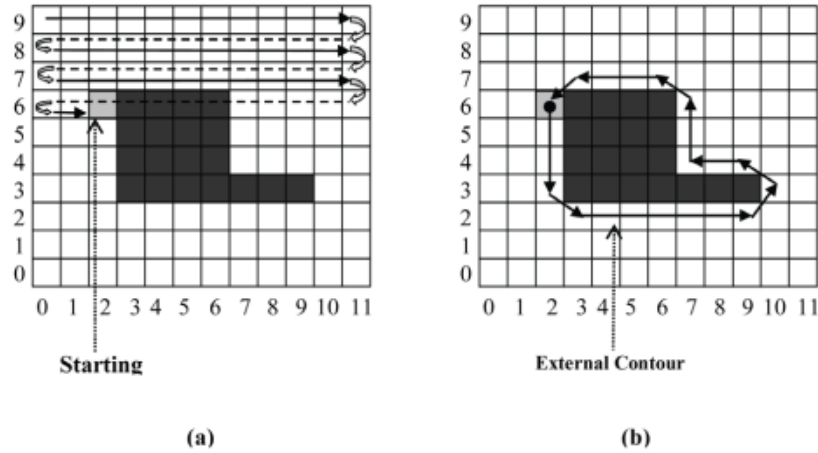
3.1.4 Extração de contorno

A partir da imagem pré-processada, o contorno de um objeto binário pode ser obtido por métodos de extração de contornos. Uma maneira importante de representar um contorno é de forma parametrizada, análoga à representação de curvas em geometria diferencial (SZELISKI, 2011).

Um algoritmo utilizado para esta extração, definido em Costa e Jr (2009), é o algoritmo de *contour following*, que acompanha a borda do objeto a partir de um ponto inicial, analisando se os pontos vizinhos também pertencem ao objeto em questão. Os passos inicial (a) e final (b) deste algoritmo podem ser visualizados na figura 2.

Este método tem como entrada uma imagem binária e, como saída, um vetor de coordenadas, que representa o contorno externo do objeto.

Figura 2 – Representação do algoritmo de extração de contorno



Fonte: Costa e Jr (2009).

3.2 Cálculo de curvatura

De acordo com Oliveira e Marroquim (2016), curvatura é a taxa de mudança na direção da aresta, em que as curvas estão presentes nos contornos (definidos por arestas).

A curvatura de uma curva regular parametrizada por uma aplicação $t \rightarrow (x(t), y(t))$, onde $x(t)$ e $y(t)$ são funções de classe C^2 , é dada por:

$$\kappa(t) = \frac{x'(t)y''(t) - y'(t)x''(t)}{(x'(t)^2 + y'(t)^2)^{3/2}}$$

Porém, caso a curva seja extraída diretamente da imagem, usamos a curvatura discreta da curva, obtida como descrito a seguir. A partir de detecção de contornos, pode-se utilizar operações vetoriais, como a diferença da direção de arestas entre pixels vizinhos da curva do modo

$$\kappa(t) = \varphi_{t+1} - \varphi_{t-1},$$

em que φ_t é a direção do gradiente do pixel t da curva. Alternativamente, pode-se calcular a direção do gradiente do pixel diretamente pela posição dos pixels:

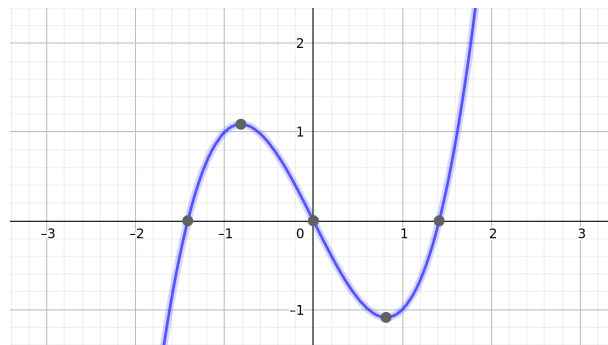
$$\kappa(t) = \frac{y_{t-1} - y_{t+1}}{x_{t-1} - x_{t+1}}$$

3.3 Características robustas

A menos de translação e rotação, uma curva pode ser determinada por sua curvatura. Considerando o gráfico da curva, os pontos de destaque são as intersecções com o eixo-x e os pontos de máximos e mínimos (locais e globais) - ver figura 3.

Os pontos de intersecção com o eixo-x representam os pontos de inflexão da curva, e os pontos extremos representam os vértices da curva. Tais pontos são denominados características robustas da curva. Observe que, pelo teorema dos quatro vértices (CARMO, 2014), toda curva fechada e simples tem, pelo menos, quatro vértices.

Figura 3 – Demarcação dos pontos importantes de uma curva com $\kappa(x) = x^3 - 2x$



Fonte: Elaborada pelo autor.

As características robustas podem ser determinadas em curvas a partir de cálculos de curvatura e da sua derivada.

3.4 Reconstrução de curvas

A reconstrução de curvas a partir de imagens ou pontos de controle é de fundamental importância em áreas como computação gráfica, geometria computacional, visão

computacional e processamento de imagens. Boa parte de soluções neste campo baseiam-se em representações do tipo *spline* (SCHOENBERG, 1946). Técnicas baseadas em *spline* consistem basicamente de um processo de minimização da distância entre a curva de referência e o resultado obtido e são ditas interpolatórias, uma vez que a curva resultante passa por todos os pontos de controle. Jakóbczak (2014) mostra como interpolar uma curva 2D a partir de um conjunto de pontos usando informação sobre a distribuição probabilística dos pontos faltantes. Mais recentemente (GAO et al., 2019) o uso de aprendizado profundo foi explorado juntamente com splines para a reconstrução de curvas paramétricas e superfícies.

Uma segunda categoria de reconstrução de curvas são as baseadas em aproximação em que não necessariamente a curva ou superfície é reconstruída passando pelos pontos de controle. Lee (2000) propõe um método baseado em mínimos quadrados combinado à uma árvore geradora mínima Euclidiana para, a partir de um conjunto de pontos não organizados, aproximar uma curva.

Uma forma alternativa para reconstruir curvas a partir de um conjunto de pontos é descrita por Sorkine (2006). A abordagem é baseada no operador de Laplace e representações diferenciais entre os vértices de uma dada vizinhança em uma malha de pontos. Em contraste com tradicional representação por coordenadas cartesianas globais, a representação diferencial de uma superfície revela informações sobre a forma local da mesma, bem como o tamanho e orientação de detalhes locais. Além de prover informações que resultarão em uma reconstrução com maior preservação de detalhes, trata-se de um sistema linear o que o torna computacionalmente eficiente.

3.4.1 O Operador Laplaciano e a representação diferencial

Seja uma malha triangular caracterizada por $M = (V, E, F)$, contendo V vértices, E arestas e F faces. Cada vértice i de M possui uma representação cartesiana dada por $\mathbf{v}_i = (x_i, y_i, z_i)$. Coordenadas diferenciais (também conhecidas como δ – *coordenadas*) de \mathbf{v}_i são definidas como a diferença entre a coordenada Cartesiana e o centro de massa de seus vizinhos imediatos na malha:

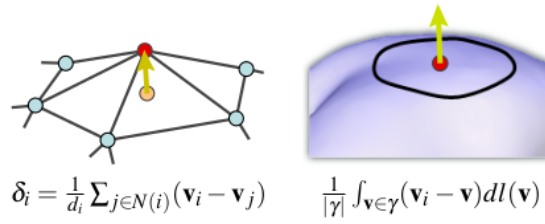
$$\delta_i = (\delta_i^x, \delta_i^y, \delta_i^z) = \mathbf{v}_i - \frac{1}{d_i} \sum_{j \in N(i)} \mathbf{v}_j, \quad (3.1)$$

onde $N(i) = \{j | (i, j) \in E\}$ e $d_i = |N(i)|$ é o número dos vizinhos imediatos de i .

Como ilustra a figura 4, as coordenadas diferenciais aproxima não apenas as características da forma local da superfície, mas também a direção normal e a curvatura média.

A equação 3.1 pode ser também representada em forma matricial, o que simplifica

Figura 4 – O vetor da coordenada diferencial em um vértice aproxima a forma local superfície: representação da direção normal e da curvatura média



Fonte: Sorkine (2006).

os cálculos. Seja a matriz de adjacência como definida pela equação 3.2 e a matriz diagonal D , com $D_{ii} = d_i$:

$$A_{i,j} = \begin{cases} 1, & (i, j) \in E \\ 0, & \text{caso contrário.} \end{cases} \quad (3.2)$$

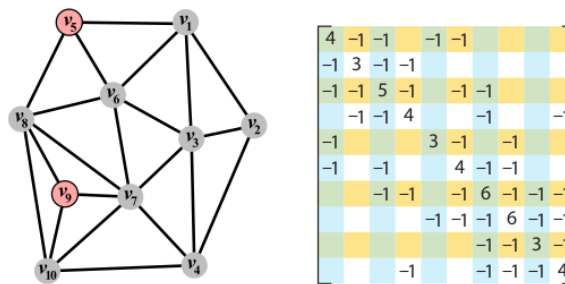
A matriz Laplaciana pode então ser calculada pela equação 3.3 ou 3.4, que é uma versão simétrica. A figura 5 ilustra um exemplo de uma simples malha triangular e sua respectiva representação matricial Laplaciana.

$$L_s = D - A \quad (3.3)$$

ou mesmo

$$(L_s)_{ij} = \begin{cases} d_i, & (i = j) \\ -1, & (i, j) \in E \\ 0, & \text{caso contrário.} \end{cases} \quad (3.4)$$

Figura 5 – Uma malha triangular e sua respectiva matriz Laplaciana L_s



Fonte: Sorkine (2006).

Uma melhor aproximação para as coordenadas diferenciais foi empregada por (PINKALL; POLTHIER, 1993) ao considerar pesos não uniformes para os vértices vizinhos.

3.4.2 Reconstrução de curvas a partir das δ -coordenadas

A princípio, dado **um conjunto** de δ - *coordenadas*, não é possível restaurar as coordenadas globais originais, pois a matriz L_s é singular, ou seja, a expressão $\mathbf{x} = L^{-1}\delta^{(x)}$ é indefinida.

No entanto, existe solução para este problema, por meio da resolução de um sistema linear do tipo **full-rank** dado por

$$\tilde{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left(||L\mathbf{x} - \delta^{(x)}||^2 + \sum_{j \in C} \omega^2 |x_j - c_j|^2 \right) \quad (3.5)$$

onde \mathbf{x} é um vetor de dimensão n que contém as coordenadas absolutas x de todos os vértices; C é o conjunto dos índices dos vértices cuja localização espacial é conhecida - o conjunto de vértices em C são restrições impostas ao sistema - também conhecidas como âncoras e $\omega > 0$ é o peso que regula a importância atribuída a posição de cada restrição. Sorkine (2006) detalha um método eficiente para se resolver o problema de mínimos quadrados da equação 3.5.

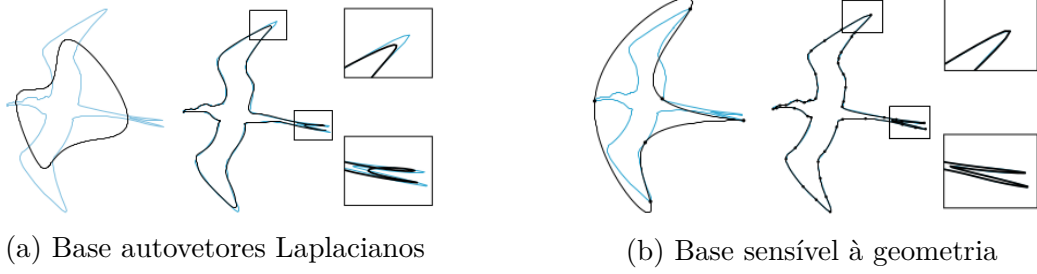
A matriz L_s possui propriedades espectrais que são fundamentais para a reconstrução de curvas a partir de um conjunto de pontos. Por ser simétrica semi-definida positiva, esta possui uma base de autovetores ortogonal $E = \{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$. Sejam os autovalores representados por λ_i , $0 = \lambda_1 < \lambda_2 \leq \lambda_3 \leq \dots \leq \lambda_n$. Sabe-se que os menores autovalores não zero λ_i são muito pequenos e o maior autovalor λ_n está associado ao vértice de maior valência na malha M . Os primeiros autovetores (associados aos menores autovalores) são suaves, enquanto os maiores têm mais altas frequências. A base Laplaciana de autovalores e autovetores é, de fato, uma extensão da base discreta de cosseno para domínios irregulares (TAUBIN, 1995). Portanto, pode-se dizer que autovalores são considerados frequências de uma malha.

Encontrar uma boa base para a representação compacta de uma forma, ou curva, significa que é possível usar somente uma fração de tal base para se aproximar de forma apropriada uma determinada geometria. Em processamento de imagens, por exemplo, pode-se utilizar bases de wavelet ou Fourier. A figura 6 ilustra um exemplo de reconstrução por meio de duas bases: a base de autovetores Laplaciana e uma base sensível à geometria (SORKINE et al., 2005). Repare que segunda base provê uma melhor aproximação, tanto em maior quanto em menor escala.

3.5 Métricas para avaliação de similaridades entre contornos

A similaridade entre curvas paramétricas, cuja representação consiste apenas de um vetor de pontos, pode ser computada simplesmente por meio de métodos tradicionais

Figura 6 – Reconstrução de uma curva fechada usando bases distintas. A reconstrução é representada pela curva negra e a curva original pelo traçado em azul. As âncoras são representadas pelos pontos negros: a) base autovetores Laplaciana 6 e 45 vértices, respectivamente; b) base sensível à geometria, também com 6 e 45 vértices, respectivamente



Fonte: Sorkine (2006).

de distância entre cada par de pontos (tanto da curva de referência quanto da curva reconstruída), como por exemplo a Distância Euclidiana ou Distância de Manhattan. No entanto, há diversas outras métricas que podem ser usadas para avaliar quantitativamente a similaridade entre dois contornos, largamente utilizadas em aplicações de visão computacional e processamento de imagens. A seguir, serão apresentadas 3 métricas: uma baseada na cobertura de regiões; a segunda, na distância entre os contornos de objetos e a terceira, uma medida baseada em probabilidades.

3.5.1 Covering

Seja S_T o resultado da segmentação e S_G uma segmentação de referência (padrão outro). A cobertura de regiões (*covering*) (CREVIER, 2008) quantifica a cobertura das regiões de S_T em relação às regiões de S_G . Esta é definida por:

$$Covering(S_T \rightarrow S_G) = \frac{1}{A} \sum_{R \in S_G} |R| \max_{R' \in S_T} \{\mathcal{O}(R, R')\}, \quad (3.6)$$

$$\mathcal{O}(R, R') = \frac{|R \cap R'|}{|R \cup R'|} = \frac{|R \cap R'|}{|R| + |R'| - |R \cap R'|}, \quad (3.7)$$

onde A é o número de pixels na imagem; R e R' são as regiões de S_G e S_T , respectivamente; $\mathcal{O}(R, R')$ representa a sobreposição entre as regiões de R e R' e $|R|$ e $|R'|$ o número de pixels nas regiões de R e R' , respectivamente. Eq. 3.6 retorna valores na faixa $(0, 1)$. Quanto maior o valor, maior a similaridade entre S_T e S_G . Esta métrica pode ser usada no contexto deste projeto, em especial em casos de curvas fechadas.

3.5.2 Medida Baseada em contornos

A avaliação de similaridade por contornos (HUANG; DOM, 1995) pode ser quantificada ao calcular-se a distância mínima entre pares de pontos de 2 conjuntos distintos: (i) BT , os contornos produzidos por um método de segmentação S_T e (ii) BG , os contornos de um padrão ouro S_G . A similaridade resultante vem do cálculo das medidas de Precisão (P), Revocação (R) e de BF1-Score, dadas por:

$$P = \frac{1}{|BT|} \sum_{p \in BT} [Matched(p, BG) \leq \theta] \quad (3.8)$$

$$R = \frac{1}{|BG|} \sum_{p \in BG} [Matched(p, BT) \leq \theta] \quad (3.9)$$

$$BF1 - Score = 2 \times \frac{P \times R}{R + P}, \quad (3.10)$$

A função $Matched(.)$, em P , percorre os pontos p de BT em busca dos pontos mais próximos dos contornos em BG , de acordo com a distância máxima θ . Se $[.]$ for verdade, esta retorna 1, caso contrário, 0. A função $Matched(.)$, em R , percorre os pontos p de BG em busca dos pontos próximos aos contornos de BT . $|BT|$ e $|BG|$ são os números de pontos existentes em BT e BG , respectivamente.

3.5.3 Probabilistic Rand Index (PRI)

A métrica PRI (UNNIKRISHNAN; HEBERT, 2005) computa a probabilidade de um par de pixels (i, j) pertencer aos contornos resultantes de um método de segmentação S_T , contendo rótulos consistentes em um conjunto k de um método de segmentação de referência S_{G_k} (padrão ouro). A métrica PRI é definida por:

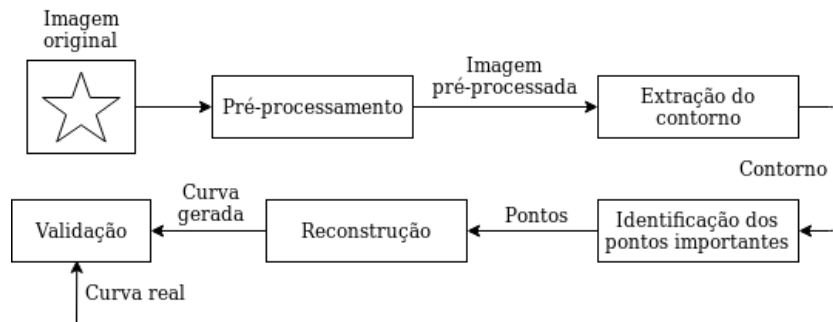
$$PRI(S_T, \{S_{G_k}\}) = \frac{1}{A} \sum_{i < j} [c_{ij}p_{ij} + (1 - c_{ij})(1 - p_{ij})] \quad (3.11)$$

onde c_{ij} é a probabilidade dos pixels i e j possuírem o mesmo rótulo na segmentação S_T e p_{ij} corresponde à probabilidade dos pixels i e j compartilharem os mesmos rótulos na segmentação de referência S_{G_k} , e A é o total do número de pares de pixels. A função PRI retorna valores na faixa $(0, 1)$. Quanto maior o valor, maior a similaridade entre S_T e S_{G_k} .

4 Materiais e métodos

O diagrama de blocos da Figura 7 ilustra as principais etapas de desenvolvimento deste projeto de pesquisa.

Figura 7 – Diagrama de bloco das etapas de desenvolvimento



Fonte: Elaborada pelo autor.

Neste projeto, optou-se inicialmente por empregar um banco de dados de folhas do *ImageCLEF 2011* (Figura 8). Seu uso se justifica pela necessidade de avaliarmos o método com precisão. Ao observarmos tais imagens é possível perceber que os contornos podem ser extraídos com uma certa facilidade e com alta precisão. Estes contornos serão utilizados como padrão ouro no processo de validação do método.

Figura 8 – Exemplo de folhas do banco de dados utilizado



Fonte: ImageCLEF (2011).

O pré-processamento consistirá em simplificar as imagens a partir de métodos de binarização, segmentação e remoção de ruídos, de modo a deixá-las preparadas para as etapas seguintes. A morfologia matemática poderá, se necessário, ser empregada na imagem binária para eliminar pequenos “buracos” no interior da imagem ou pontos espúrios ao longo da borda.

Será então obtido o contorno da imagem pré-processada, que será representado como curvas discretas. Nestas serão feitas cálculos de variação de curvatura, de modo a encontrar seus pontos importantes.

A partir dos pontos importantes (ou características robustas) será feita a reconstrução das curvas.

Com as curvas originais extraídas das imagens e as curvas geradas pelo algoritmo de reconstrução a partir de pontos importantes, os resultados podem ser avaliados por

meio de métodos tradicionais de distância entre pontos (entre os pontos corretos das curvas e os pontos gerados pelo algoritmo) ou por meio das métricas descritas na seção 3.5.

Vale lembrar que esta proposta é parte de um projeto temático que visa a reconstrução de elementos de uma face humana. Propomos também avaliar os métodos implementados com base em dados reais de faces e caricaturas humanas (como exemplo na figura 9). A diferença é que neste tipo de imagem teremos não apenas um contorno, mas uma coleção destes.

Figura 9 – Caricatura humana



Fonte: Russell (2011).

O projeto será implementado em plataforma Linux, com as linguagens de programação *MatLab* e, caso necessário, *Python*, utilizando-se bibliotecas padrões destas linguagens. Também poderá ser utilizada a linguagem *C++*, caso seja necessária uma implementação com menor tempo de execução.

5 Plano de trabalho e cronograma

As ações definidas neste projeto de iniciação científica serão executadas de acordo com o cronograma definido na tabela 1. Como o aluno já tem noções de processamento de imagens, acreditamos que as etapas de pré-processamento e extração de contornos será feita em menor tempo. A etapa que demandará mais esforço, que inclui não apenas a implementação mas o estudo e compreensão do método de Sorkine (2006), será a reconstrução da curva a partir dos pontos importantes.

Durante todo o andamento do projeto serão realizadas reuniões periódicas com professores - tanto o orientador quanto outros professores da equipe do projeto temático - de modo a complementar e auxiliar a pesquisa. Vale ressaltar que o candidato tem participado regularmente de tais reuniões (atualmente em modo sempre virtual) que acontecem semanalmente.

Atividades	Meses de trabalho					
	1º e 2º	3º e 4º	5º e 6º	7º e 8º	9º e 10º	11º e 12º
Estudo das técnicas de reconstrução de curvas	•	•				
Pré-processamento e Extração de contornos		•	•			
Extração dos pontos importantes		•	•			
Redação do relatório parcial			•			
Implementação da reconstrução de curvas (Sorkine)			•	•	•	
Avaliação e testes				•	•	
Desenvolvimento do relatório final					•	•

Tabela 1 – Cronograma de atividades para 12 meses de trabalho.

Referências

- BERG, M. de et al. *Computational Geometry: Algorithms and applications*. 2. ed. Berlin, Heidelberg: Springer, 2000. 183-200 p.
- BORROTO-FERNÁNDEZ, M.; MECIAS, A. L.; HIDALGO, M. G. Image smoothing using the perona-malik equation with a new estimation method of the contrast parameter. 10 2013.
- BRUCE, J. W.; GIBLIN, P. J. *Curves and Singularities: A Geometrical Introduction to Singularity Theory*. 2. ed. Cambridge: Cambridge University Press, 1992.
- BURGER, W.; BURGE, M. J. *Digital Image Processing: An algorithmic introduction using java*. 2. ed. Londres: Springer, 2016. 253-263 p.
- CARMO, M. P. do. *Geometria Diferencial de Curvas e Superfícies*. São Paulo: SBM, 2014.
- COSTA, L. da F.; JR, R. M. C. *Shape Classification and Analysis*. 2. ed. Boca Raton: CRC Press, 2009. 4, 323-335 p.
- CREVIER, D. Image segmentation algorithm development using ground truth image data sets. *Comput. Vis. Image Underst.*, Elsevier Science Inc., New York, NY, USA, v. 112, n. 2, p. 143–159, nov. 2008. ISSN 1077-3142. Disponível em: <<http://dx.doi.org/10.1016/j.cviu.2008.02.002>>.
- GAO, J. et al. Deepspine: Data-driven reconstruction of parametric curves and surfaces. *CoRR*, abs/1901.03781, 2019. Disponível em: <<http://dblp.uni-trier.de/db/journals/corr/corr1901.html#abs-1901-03781>>.
- GONZALEZ, R. C.; WOODS, R. E. *Digital Image Processing*. 4. ed. Nova Iorque: Pearson, 2018. 761-826 p.
- HUANG, Q.; DOM, B. Quantitative methods of evaluating image segmentation. In: *Proceedings., International Conference on Image Processing*. [S.l.: s.n.], 1995. v. 3, p. 53–56 vol.3.
- IMAGECLEF. 2011. Disponível em: <<https://www.imageclef.org/2011/Plants>>. Acesso em: 01 may 2020.
- IZUMIYA, S. et al. *Differential Geometry From A Singularity Theory Viewpoint*. Singapura: World Scientific, 2015. 139 p.
- JAKÓBCZAK, D. J. Curve interpolation and shape modeling via probabilistic nodes combination. *Vietnam Journal of Computer Science*, Springer Science and Business Media LLC, v. 1, n. 3, p. 141–153, abr. 2014. Disponível em: <<https://doi.org/10.1007/s40595-014-0016-7>>.
- LEE, I.-K. Curve reconstruction from unorganized points. *Computer Aided Geometric Design*, Elsevier BV, v. 17, n. 2, p. 161–177, fev. 2000. Disponível em: <[https://doi.org/10.1016/s0167-8396\(99\)00044-8](https://doi.org/10.1016/s0167-8396(99)00044-8)>.

- OLIVEIRA, A.; MARROQUIM, R. *Processamento de Imagens*: aula 09 - curvatura. 2016. Disponível em: <<https://www.lcg.ufrj.br/marroquim/courses/cos756/lessons/09-curvature.pdf>>. Acesso em: 7 mar 2020.
- OTSU, N. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, v. 9, n. 1, p. 62–66, 1979.
- PERONA, P.; MALIK, J. Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 12, n. 7, p. 629–639, 1990.
- PINKALL, U.; POLTHIER, K. Computing discrete minimal surfaces and their conjugates. *Experimental Mathematics*, Informa UK Limited, v. 2, n. 1, p. 15–36, jan. 1993. Disponível em: <<https://doi.org/10.1080/10586458.1993.10504266>>.
- REGAZZONI, C. S.; FORESTI, G. L.; VENETSANOPOULOS, A. N. *Mathematical Morphology and Its Applications to Image Processing*. Fontainebleau: Springer, 1994. v. 2. 186 p.
- RUSSELL, E. *Fundamentals of Caricature Drawing in Photoshop*. 2011. Disponível em: <<https://www.pluralsight.com/courses/fundamentals-caricature-drawing-photoshop-66>>. Acesso em: 05 may 2020.
- SCHOENBERG, I. J. Contributions to the problem of approximation of equidistant data by analytic functions. part a. on the problem of smoothing or graduation. a first class of analytic approximation formulae. *Quarterly of Applied Mathematics*, American Mathematical Society (AMS), v. 4, n. 1, p. 45–99, abr. 1946. Disponível em: <<https://doi.org/10.1090/qam/15914>>.
- SERRA, J. *Image Analysis and Mathematical Morphology*. Orlando: Academic Press, 1984. ISBN 9780126372427.
- SORKINE, O. Differential representations for mesh processing. *Computer Graphics Forum*, European Association for Computer Graphics, v. 25, n. 4, p. 789–807, 2006.
- SORKINE, O. et al. Geometry-aware bases for shape approximation. *IEEE Transactions on Visualization and Computer Graphics*, Institute of Electrical and Electronics Engineers (IEEE), v. 11, n. 2, p. 171–180, mar. 2005. Disponível em: <<https://doi.org/10.1109/tvcg.2005.33>>.
- SZELISKI, R. *Computer Vision: Algorithms and applications*. London: Springer, 2011.
- TAUBIN, G. A signal processing approach to fair surface design. In: *Proceedings of the 22nd annual conference on Computer graphics and interactive techniques - SIGGRAPH 95*. ACM Press, 1995. Disponível em: <<https://doi.org/10.1145/218380.218473>>.
- UNNIKRISHNAN, R.; HEBERT, M. Measures of similarity. In: *2005 Seventh IEEE Workshops on Applications of Computer Vision (WACV/MOTION'05) - Volume 1*. [S.l.: s.n.], 2005. v. 1, p. 394–394. ISSN null.