1 ABORDAGEM BASEADA EM CURVAS

Neste capítulo é apresentada uma visão geral das recentes abordagens e o modelo multi-estágios proposto por Guo et al. (2014a), Guo, Narayanan e Kimia (2017) que é a base para a especialização construída neste trabalho. Constituída por vários estágios, parte da aquisição de elementos de borda e vai até a extração final de curvas, utilizando agrupamentos locais e globais. São usadas técnicas de aprendizado não-supervisionado em grafos obedecendo a uma coerência geométrica, seguidas pelo aprendizado supervisionado para o agrupamento e seleção global utilizando os agrupamentos anteriores. No próximo capítulo são apresentados detalhes técnicos dos métodos aqui descritos.

1.1 Terminologia

Ao longo do texto é usada uma terminologia desenvolvida para identificar as estruturas de dados utilizadas. São listados aqui alguns dos termos detalhando-os ao longo do texto conforme necessário:

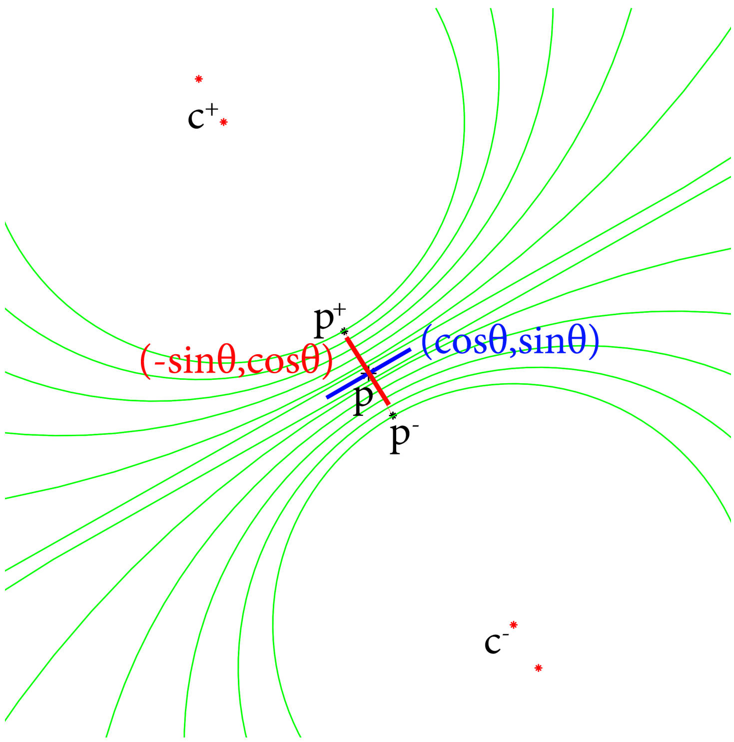
- *edgel* — advém da junção de *edge* e *pixel* que, em português, é denominado simplesmente de elemento de "borda". Sendo uma estrutura básica que representa o menor fragmento de uma borda ou contorno, o *edgel* é constituído por um ponto 2D, com direção e contraste, definido como $\boldsymbol{e} = (p, \theta, \Delta)$, com $\boldsymbol{p} = (x, y) \in \mathbb{R}^2$, θ o ângulo da tangente ao contorno no ponto p e Δ o módulo do gradiente da intensidade do sinal (contraste);

- *edgel chain* — um *edgel chain*, aqui também chamado de fragmento de curva ou apenas curva (discreta), é uma sequência ordenada de *edgels*, obedecendo critérios de continuidade e consistência geométrica, sendo a representação final de uma curva que tenha sido extraída de uma imagem;

- *edgemap* — dada uma grade 2D definida sobre a imagem, o *edgemap* ou mapa de bordas atribui a cada par (i, j) da grade o conjunto de *edgels* associado àquele píxel. Mais especificamente, o *edgemap* pode ser representado por uma matriz de células, EM, onde cada célula ou é vazia, $(EM)_{i,j} = \emptyset$, ou contém os correspondentes *edgels*, $(EM)_{i,j} = \{\boldsymbol{e}_0, ..., \boldsymbol{e}_k\}$, onde os $\boldsymbol{e}_i$'s são *edgels* do píxel (i, j). Esta estrutura dá acesso direto aos *edgels* a partir de sua posição 2D discreta;

- *curve bundle* — uma *curve bundle* ou feixe de curvas é uma estrutura contendo um conjunto de curvas paramétricas (representados por um modelo de curva e os parâmetros desse modelo) passando por perturbações de um ou mais *edgels*. Por

Figura 5 - Exemplo de construção das curvas representadas por um *curve bundle*



Legenda: Exemplo de construção das curvas representadas por um *curve bundle*. Dado um *edgel* $\boldsymbol{e} = (\boldsymbol{p}, \theta, \Delta)$, e uma curvatura k, constrói-se pares de semi-círculos definidos pelos parâmetros $(\boldsymbol{p}^*, \theta^*, k^*)$, com $\boldsymbol{p}^* \in [\boldsymbol{p} - \delta n, \boldsymbol{p} + \delta n]$, $\theta^* \in [\theta - \delta\theta, \theta + \delta\theta]$ e $k^* \in [k - \delta k, k + \delta k]$, para δn, $\delta\theta$ e δk em intervalos previamente especificados.

Fonte: O autor, 2017.

exemplo, um arco circular do feixe pode ser representado por $(\delta n, \delta\theta, \delta k)$, onde esses parâmetros são relativos ao *edgel* âncora do feixe — o *edgel* a partir do qual se iniciam as curvas desse feixe — sendo δn uma perturbação na posição $\boldsymbol{p}$, $\delta\theta$ uma perturbação na orientação θ do *edgel* âncora deste feixe, e δk uma perturbação na curvatura k do arco circular, todas dentro de um intervalo restrito. Embora a posição $\boldsymbol{p}$ seja 2D, a perturbação na posição δn é 1D sendo esta perpendicular à tangente no ponto $\boldsymbol{p}$ (direção normal);

- *curvelet*, conjunto ordenado de *edgels* $E = \{\boldsymbol{e}_0, ..., \boldsymbol{e}_k\}$ e um feixe de curvas $f = (\delta n, \delta\theta, \delta k)$ referente a um *edgel* âncora $\boldsymbol{e}_i \in E$, sendo que este feixe de curvas é coerente com todos os *edgels* em E.

Exemplo de construção de um par de curvas de uma *curve bundle* com modelo de curva circular. Seja dado um *edgel* $\boldsymbol{e} = (p, \theta, \Delta)$ e a curvatura k. O vetor unitário tangente ao *edgel* é $(\cos\theta, \sin\theta)$, $1/k$ é o raio dos semi-círculos tangentes ao *edgel*, e seus centros são

$$\boldsymbol{c}^{+} = \boldsymbol{p} + \frac{1}{k}(-\sin\theta, \cos\theta),$$
$$\boldsymbol{c}^{-} = \boldsymbol{p} - \frac{1}{k}(-\sin\theta, \cos\theta).$$

Finalmente os semi-círculos são dados parametricamente por

$$C^{+} = \{\boldsymbol{c}^{+} + \frac{1}{k}(\cos\alpha, \sin\alpha), \quad \forall\alpha \in [\theta - \pi, \theta]\},$$
$$C^{-} = \{\boldsymbol{c}^{-} + \frac{1}{k}(\cos\alpha, \sin\alpha), \quad \forall\alpha \in [\theta, \pi + \theta]\}.$$

Na Figura 5 são ilustrados alguns semi-círculos com variações na posição $\boldsymbol{p}$ e curvatura k, mantém-se θ fixo apenas para uma melhor visualização, sem curvas sobrepostas.

1.1.1 Critérios para avaliação dos fragmentos de curvas

Se fragmentos de curva forem considerados como uma representação intermediária dos contornos suaves detectados em uma imagem, as características que tornam o seu conjunto semanticamente consistente podem ser explicitadas pelas propriedades semânticas:

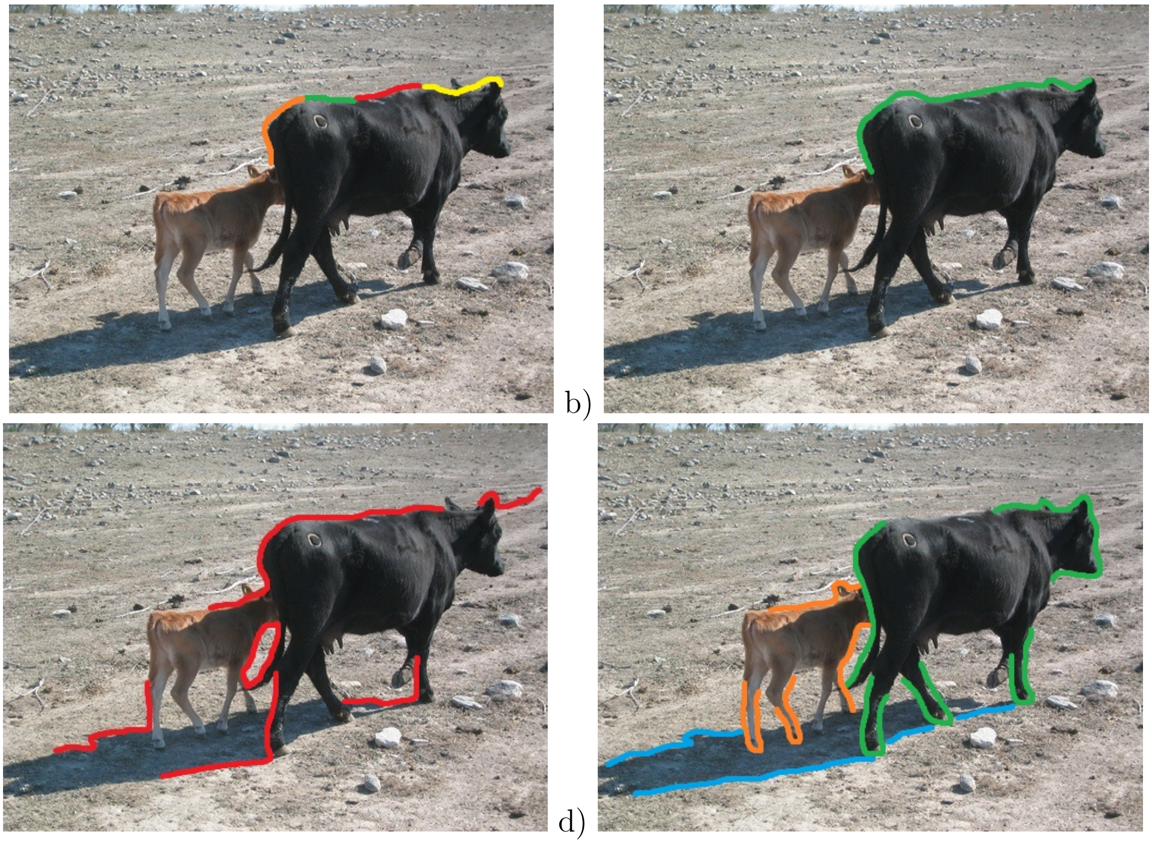
- **totalidade** — fragmentos pertencentes a um mesmo contorno são agrupados;
- **distinção** — fragmentos pertencentes a contornos distintos são separados;
- **objetividade** — todos os fragmentos semanticamente válidos pertencem ao conjunto (são extraídos se existem evidências disso).

A totalidade e distinção estão relacionadas ao agrupamento dos elementos de borda em curvas enquanto a objetividade refere-se ao conjunto de contornos verídicos. Assim a solução ótima deve extrair contornos geometricamente suaves obedecendo tais propriedades, como ilustrado na Figura 6.

Os contornos extraídos das imagens são referenciados ao longo desta dissertação também como:

- verídico — contorno com evidências na imagem provenientes dos fatores físicos mencionados na Figura 3;
- não-verídico — também chamado de alucinado, é aquele que apresenta evidências na imagem mas essas são provenientes de ruídos como, por exemplo, compressão;

Figura 6 - Requerimentos na extração de curvas

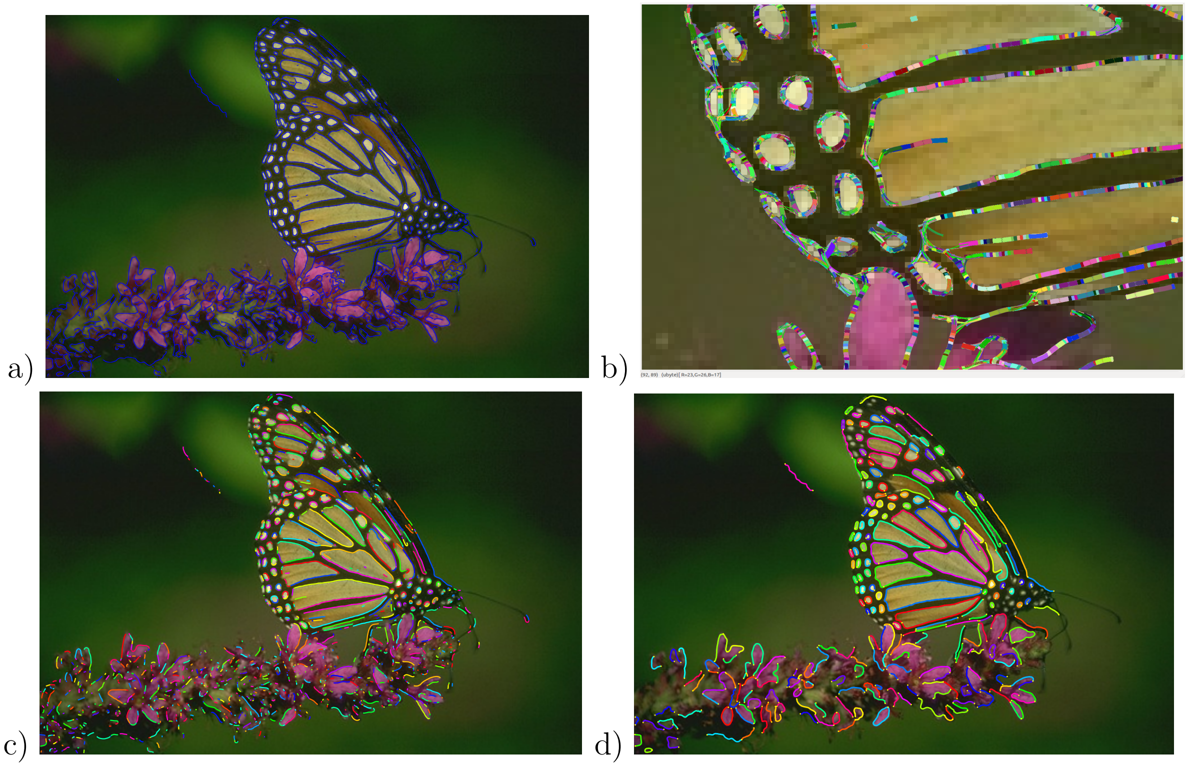


a) b) c) d)

Legenda: Requerimentos na extração de curvas.
(a) e (b) representam respectivamente como contornos extraídos em (a) deveriam ser agrupados em (b) relativamente à totalidade.
(c) e (d) exibem respectivamente, curvas sem distinção, porém objetivas, em (c) e com distinção e objetividade em (d).

Fonte: (GUO; NARAYANAN; KIMIA, 2017).

Figura 7 - Visão geral dos principais estágios na extração genérica de curvas



Legenda: Visão geral dos principais estágios na extração genérica de curvas.
a) Bordas detectadas em azul; estas bordas ainda não estão agrupadas em curvas.
b) *Curvelets*, pequenas curvas que localmente obedecem ao modelo de curva de arco circular, construídas a partir dos *edgels* locais.
c) Fragmentos de curvas finais extraídos com o método de Tamrakar e Kimia (2007).
d) Fragmentos de curvas extraídos com o novo estágio proposto por Guo et al. (2014a), Guo, Narayanan e Kimia (2017).

Fonte: (GUO et al., 2014a; GUO; NARAYANAN; KIMIA, 2017)

- desejável — contorno verídico que diz respeito à geometria da cena e seja útil à aplicação em tela;
- indesejável — contorno não-verídico ou proveniente de fatores físicos que não dizem respeito à geometria da cena desejada em uma dada aplicação.

1.2 Abordagens recentes

A abordagem de Guo et al. (2014a), Guo, Narayanan e Kimia (2017) que soluciona em cascata, o bem conhecido problema de extração de curvas, como o conceito original proposto por Viola e Jones (2001), é estruturada em multi-estágios, onde a suavidade a nível de elementos de borda (*edgels*) é tratada nos estágios iniciais enquanto grupos de *edgels* conectados são tratados nos últimos estágios, gerando fragmentos de curva semanticamente consistentes, (Figura 7). Cada problema de grupamento é tratado separadamente, lidando apropriadamente com a eficiência computacional e aplicabilidade a partir de dados de treinamento reduzidos.

1.2.1 Relação com trabalhos anteriores

A necessidade de não-linearidade no nível de detecção de bordas foi identificada por Iverson e Zucker (1995), que propuseram o uso de uma função não-linear na forma de um operador lógico-linear ao invés de somas lineares sobre o suporte de um detector de elementos de borda. A abordagem de Guo et al. (2014a), Guo, Narayanan e Kimia (2017), utilizada aqui, é similar a essa ideia no nível de fragmentos de curva. A literatura de detecção de fronteiras é ampla, porém a maioria foca na detecção de bordas ou mapa de saliência, poucos se dedicam à detecção de fragmentos de curvas.

1.2.2 Extração clássica de curvas

O trabalho de Fischler, Tenenbaum e Wolf (1981) é canônico na área, tendo sido seguido nas três décadas posteriores. Um detector de estradas em imagens aéreas de baixa resolução é proposto, o qual gera um mapa espacial no qual é aplicado o algoritmo de busca A^* (lê-se "A estrela") para encontrar o caminho de custo mínimo entre dois pontos. Muitos trabalhos seguem essencialmente este padrão de busca usando uma função objetivo linear. A desvantagem desses modelos é que para qualquer caminho com um baixo custo (alta evidência de bordas) existe a premissa chave que o custo é distribuído uniformemente. Entretanto, esta hipótese não é sempre verdadeira, permitindo que existam caminhos com alta evidência mas esparsos, conduzindo a um grande número de falso-positivos, curvas alucinógenas, ou perdas na detecção.

1.2.3 Literatura recente

Grande parte da literatura existente pode ser apresentada na forma de um problema de otimização global. Seja $V = \{v_i | i = 1, ..., N\}$ o conjunto de possíveis amostras ao longo de um fragmento de curva, podendo ser píxeis, Felzenszwalb e McAllester (2006), elementos de borda na imagem (*edgels*), Zhu, Song e Shi (2007), ou pequenos agrupamentos de bordas (super-nós) na forma de *tokens* de linha, Kokkinos (2009) ou sub-grafos locais, Kennedy, Gallier e Shi (2011). Seja $E = \{e_{ij} | d(v_i, v_j) < d_0, \forall v_i, v_j \in V\}$ o conjunto de todas as possíveis conexões, entre as amostras, de todas as curvas candidatas. O grafo $G(V, E)$ é um super-conjunto dos possíveis fragmentos de curva. Assim, o problema de extração de curvas pode ser posto como a seleção de um subconjunto de conexões $E^* \subset E$, tal que:

$$E^* = \underset{\bar{E} \subseteq E}{\operatorname{argmin}} [f_{fg}(\bar{E}) + f_{bg}(E \setminus \bar{E})], \tag{1}$$

onde f_{fg} e f_{bg} são funcionais objetivo, capturando características do plano principal da curva (*fg* — *foreground*) e do plano de fundo (*bg* — *background*) respectivamente. Por exemplo, na abordagem *min-cover*, Felzenszwalb e McAllester (2006), v_i são píxeis na imagem com o algoritmo buscando minimizar um funcional objetivo como na equação 1. As duas abordagens KGS, Zhu, Song e Shi (2007), Kennedy, Gallier e Shi (2011), e FPG, Kokkinos (2009), seguem uma estratégia de dois estágios: no KGS o corte normalizado, Shi e Malik (1997), divide o grafo em aglomerados representando v_i. No FPG pequenas retas são formadas representando v_i. Ambos os métodos seguem otimizando o problema na forma da equação 1 com funcionais objetivo lineares.

A otimização em dois estágios aumenta a eficiência, mas ainda assim apresenta sérias desvantagens: a aglomeração inicial com o corte normalizado pode introduzir conexões espúrias no corte k-normalizado, Yu e Shi (2003), que inclui otimização não-convexa, inicializadores e aglomerados aleatórios resultando em pouca estabilidade. Similarmente, o estágio inicial do FPG aparenta ser excessivamente conservador, dada as limitações geométricas impostas pelo modelo linear, deixando a maior parte do trabalho para a segunda fase. Do ponto de vista da representação, a extração de curvas como um problema de otimização global, Felzenszwalb e McAllester (2006), Kokkinos (2009), Kennedy, Gallier e Shi (2011), é mais seletiva na obtenção de longos contornos semanticamente verídicos, mas negligencia pequenas curvas geometricamente verídicas por causa de questões de tratabilidade computacional, esbarrando em um limite superior de cobertura relativamente baixo comparado com algoritmos de agrupamento local.

1.3 De elementos de borda a curvas

A junção de elementos pontuais de borda em curvas se dá através de seis estágios, iniciando por considerações locais apenas geométricas (estágios I a IV, tratados nesta seção) até considerações globais semânticas (estágios V e VI, tratados na seção seguinte), mantendo a cobertura (fração entre bordas verídicas extraídas pelo algoritmo e as bordas desejadas) enquanto reduz o espaço de busca eliminando rapidamente as curvas indesejadas. Especificamente, a extração não-supervisionada "bruta" de curvas acontece nos quatro primeiros estágios, Figura 8:

I. O mapa de bordas é computado com limiar de detecção baixo, ou seja, variações mínimas do contraste no sinal da imagem são considerados elementos de borda, consequentemente com alta cobertura das bordas verídicas mas grande quantidade de falso-positivos proveniente de ruídos. A Figura 36 é um exemplo prático;

II. Formação de *curvelets* em vizinhanças de cada *edgels*. Essas *curvelets* são geometricamente consistentes, reduzindo falso-positivos geométricos com perda insignificante na cobertura;

III. *Edgels* que pertençam a uma única *curvelet*, (ou seja, existe apenas um feixe de curvas viável passando por esses *edgels*), são extraídos, deste modo formando fragmentos de curvas sem ambiguidade que serão o ponto de partida para a resolução das ambiguidades;

IV. As ambiguidades restantes são os pontos de junção entre dois ou mais fragmentos não ambíguos extraídos anteriormente. Estas ambiguidades são representadas por árvores de hipóteses com raiz na extremidade de cada fragmento de curva não-ambíguo. A resolução destas ambiguidades dá-se através da aplicação de um funcional objetivo, que leva em conta a suavidade das possíveis curvas e minimiza lacunas. Dessa forma conectam-se os fragmentos de curva restantes formando os fragmentos finais.

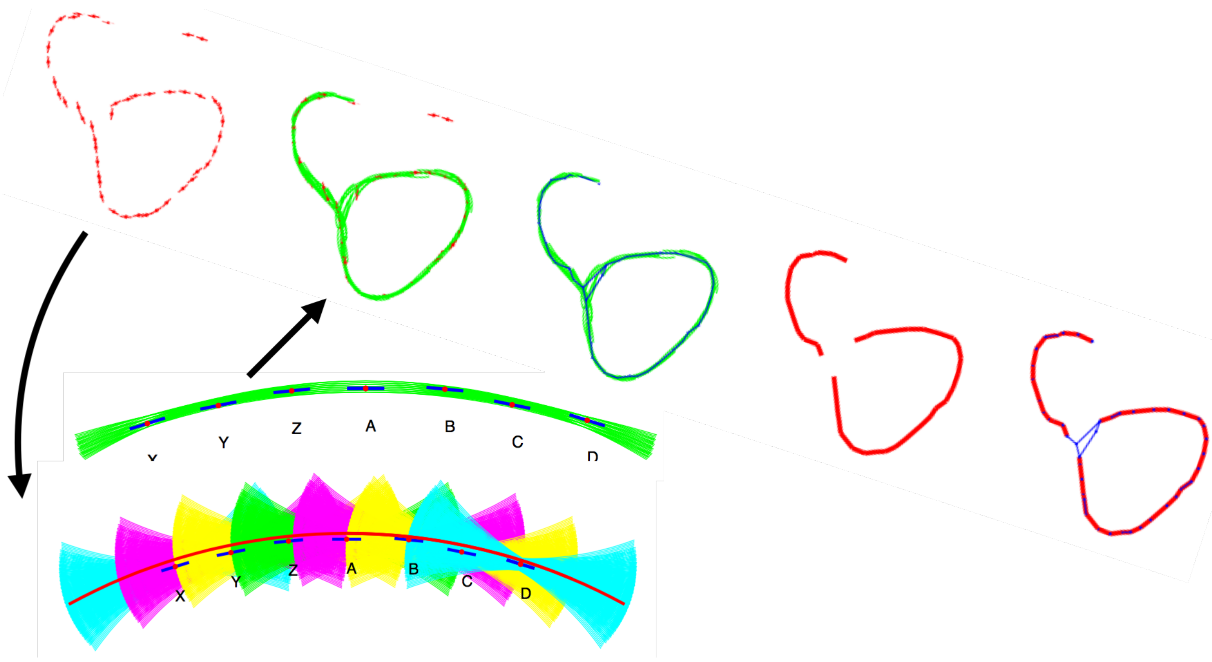
Em seguida detalha-se os quatro estágios brevemente descritos anteriormente. Exemplos mais concretos são dados no capítulo 4 e na figura 13.

1.3.1 Estágio I: detecção de bordas

Este estágio tem como entrada uma imagem, geralmente em tons de cinza, da qual se pretende extrair (determinar) as bordas, e tem como saída um mapa de elementos de borda, com os *edgels* (elementos de borda) detectados. Esses elementos de borda são detectados a partir da variação na intensidade da imagem utilizando o método de terceira ordem, Tamrakar e Kimia (2007).

A detecção de bordas tem uma extensa literatura incluindo dois do estados-da-arte mais populares *global probability boundary* (gPb), Arbelaez et al. (2011), e *structured edge* (SE), Dollár e Zitnick (2015); porém, escolheu-se usar especificamente o método de terceira ordem, baseado no gradiente e em geometria diferencial, pois este tem um cuidado especial com a orientação dos *edgels*, o que influencia diretamente os estágios seguintes, dado que a construção de curvas geometricamente consistentes está diretamente ligada às orientações das bordas. Na Figura 9 é apresentado um exemplo da qualidade das direções

Figura 8 - Vinculador simbólico de *edgels*



Legenda: Vinculador simbólico de *edgels*. Emprega-se a aglomeração geométrica, obedecendo as propriedades geométrico-diferenciais, podendo ser expressada como uma transformada local intrínseca de Hough, Tamrakar e Kimia (2007), com aprimoramentos por Guo et al. (2014b). Na parte superior são ilustrados os estágios, partindo de elementos de borda, traços vermelhos, à construção de *curvelets*, em verde, com as curvas possíveis, em azul, e finalmente as curvas sem ambiguidade em vermelho. O detalhe na parte inferior ilustra uma *curvelet* passando pelos *edgels* X a D, com seus feixes de curvas multicoloridos.

Fonte: (TAMRAKAR, 2008).

Figura 9 - Detector de elementos de borda subpíxel.

Legenda: *Edgels* obtidos pelo detector subpíxel, em verde, e com outros métodos tradicionalmente utilizados, em vermelho. É notória a acurácia nas orientações dos *edgels* em verde. A orientação é crucial nos estágios subsequentes.

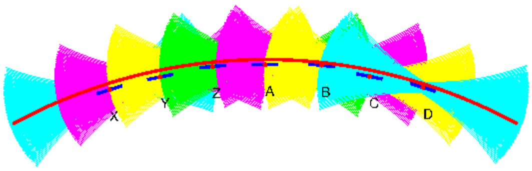
Fonte: (TAMRAKAR; KIMIA, 2007).

desse método em comparação com outros métodos que utilizam a direção da intensidade do gradiente.

1.3.2 Estágio II: agrupamento de bordas em *curvelets*

O agrupamento neste estágio é feito localmente, seguindo um modelo de curva de arco circular. Cada *edgel* do mapa de bordas, adquirido no estágio anterior, tem um feixe de curvas padrão definido pelas perturbações na posição e na orientação do *edgel* e no parâmetro do modelo de curva, que neste caso de arco circular é a curvatura. Dessa forma, o agrupamento local dos *edgels* com seus feixes de curvas, se traduz em encontrar a interseção dos espaços de parâmetros desses feixes de curva, Tamrakar (2008), Tamrakar e Kimia (2007). Esta interseção está em $\mathbb{R}^3$, definida pelos parâmetros, a perturbação na posição do *edgel*, a perturbação na orientação do *edgel* e a curvatura k do arco circular. Este tipo de problema, encontrar curvas que passam por pontos dados no espaço intrínseco de parâmetros dessas curvas, é remanescente da transformada de Hough.

Figura 10 - Vários *edgels* nomeados de X a Z e de A a D, e seus feixes de curva multi coloridos.



Legenda: Vários *edgels* nomeados de X a D, e seus feixes de curva multi coloridos. Em vermelho uma curva viável geometricamente, passando por pertubações dos *edgels*.
Fonte: (TAMRAKAR, 2008).

O processo de redução do espaço de parâmetros é feito sequencialmente: a cada novo *edgel* adicionado se restringe as possíveis curvas às que passam por todos eles, obedecendo limites impostos pelo modelo de curva de arco circular. Nas figuras 10 e 11 é exemplificada a redução do espaço das possíveis curvas.

Assim, são construídas todas as possíveis curvas locais, a partir de cada *edgel* do mapa de bordas, geometricamente consistentes com o modelo de arco circular. Essas pequenas curvas são as *curvelets* e são a saída deste estágio.

1.3.3 Estágio III: traçando fragmentos de curvas não-ambíguas

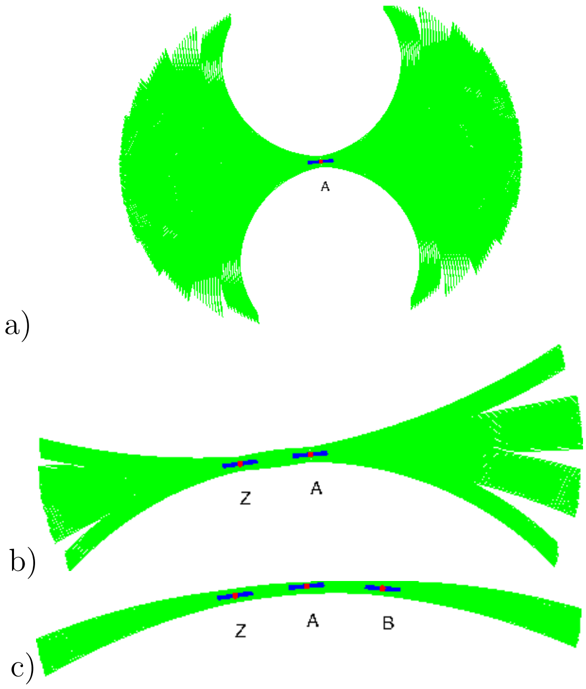
A formação de *curvelets* a partir de cada *edgel* dá origem a ambiguidades, já que todas as possíveis *curvelets* são geradas. Neste estágio extrai-se as partes não ambíguas, ou seja, os *edgels* que estão estritamente conectados a apenas um outro *edgel*. Esta extração é feita utilizando componentes conexas em um grafo de *edgels*, Tamrakar (2008), onde a conexão entre dois *edgels* quaisquer existe se, e somente se, existe alguma *curvelet* contendo esses *edgels* e eles são sucessivos nessa *curvelet*.

Esses fragmentos de curvas extraídos do grafo são utilizadas como chave na posterior resolução das ambiguidades, pois são geometricamente consistentes, logo existe ao menos uma *curvelet* com o modelo de curva de arco circular passando por quaisquer dois *edgels* consecutivos desse fragmento de curva, e provavelmente são partes de segmentos globais verdadeiros.

1.3.4 Estágio IV: resolução de ambiguidades

A abordagem utilizada na resolução de ambiguidades é a construção de árvores de hipóteses, contendo as possíveis conexões entre os segmentos não-ambíguos que contenham evidências entre suas extremidades, Figura 12. Essas árvores são construídas para cada

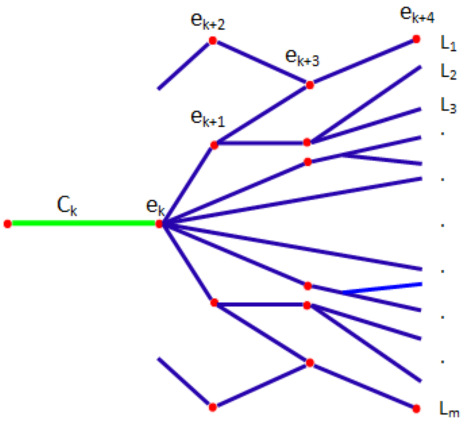
Figura 11 - Redução das possíveis curvas passando por, um, dois e três *edgels*.



Legenda: Redução das possíveis curvas passando por um (a), dois (b) e três (c) *edgels*. Muitas vezes com poucos *edgels* já não é possível ter uma curva viável, removendo qualquer possibilidade posterior de formação de *curvelets* com mais *edgels*.

Fonte: (TAMRAKAR, 2008).

Figura 12 - Árvore de hipóteses.

a) 

b) 

Legenda: Árvore de hipóteses.

a) Em verde um contorno não ambíguo com as possíveis conexões entre os *edgels* ambíguos, em azul, na forma de árvore.

b) Exemplo real de como se comportam tais hipóteses em uma imagem

Fonte: (GUO et al., 2014c).

extremidade dos segmentos não ambíguos. Cada árvore dá origem a várias conexões entre os fragmentos. As possibilidades nas extremidades de cada fragmento de curva são:

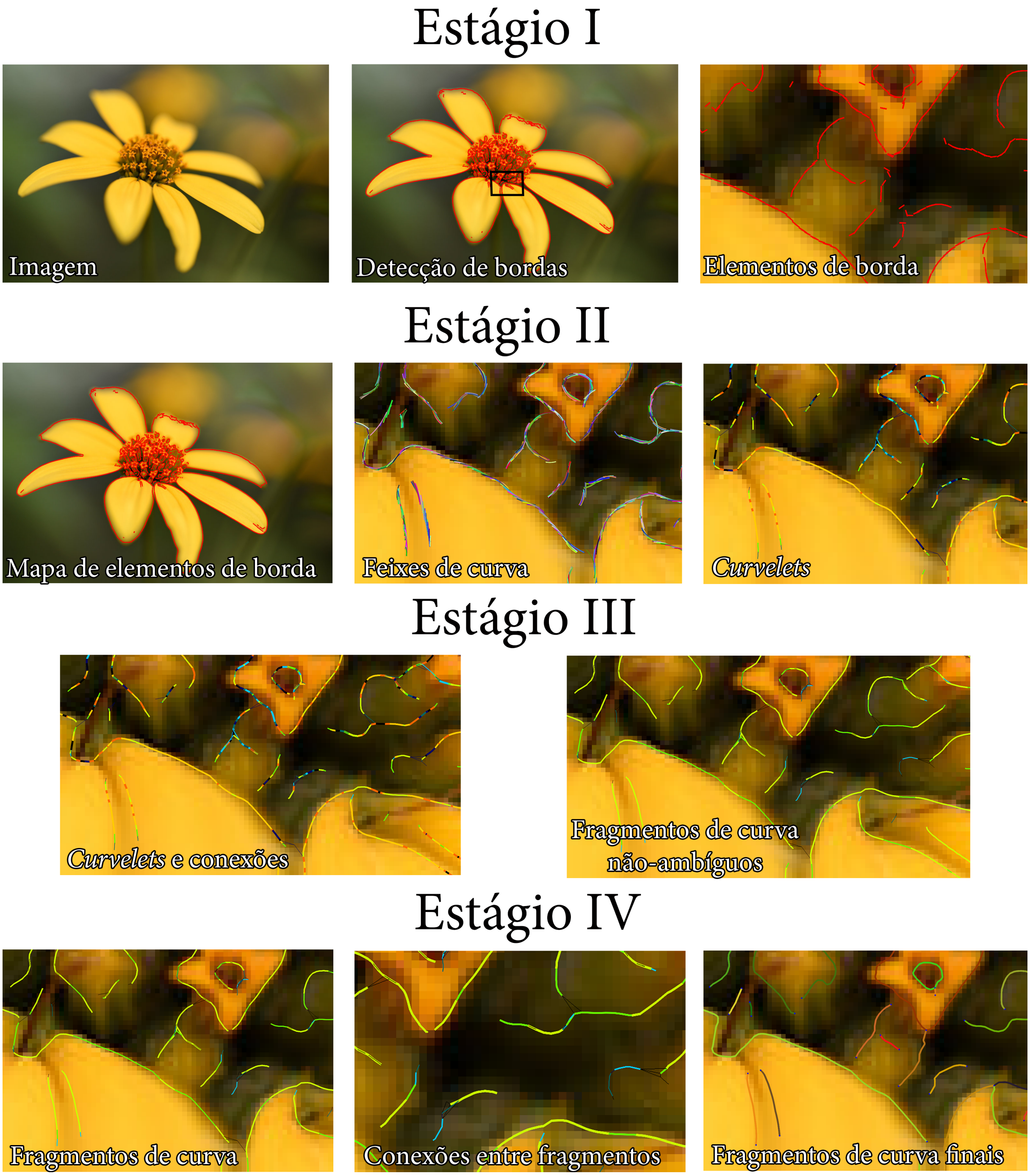
- a extremidade é realmente um ponto final do segmento;
- o fragmento se estende por outro segmento com ambiguidade em um certo ponto;
- o fragmento se estende por outros dois segmentos em uma junção “Y” ou “T”;
- o fragmento se estende mas não necessariamente termina em um outro fragmento.

A decisão de qual dos casos é o verdadeiro para cada curva possível é feita percorrendo cada árvore de hipóteses otimizando uma função objetivo como a da equação 1, com uma estratégia do tipo *best-first*. Este tipo de estratégia gulosa (*greedy*) inicia no nó raiz da árvore de hipóteses e expande o nó folha mais promissor dado pela função objetivo. Desta forma, encontra-se um caminho na árvore de hipóteses que é bom o suficiente, embora não seja a solução global. Os fragmentos de curva deste caminho são então conectados formando um único fragmento. Vale notar que a ordem de formação das árvores de hipóteses influencia no resultado, dado que este é um algoritmo guloso.

1.4 Aprendizado da topologia de curvas

Os quinto e sexto estágios são uma abordagem nova, o uso de aprendizado de máquina supervisionado para aprimorar a geometria e a semântica topológica. Os frag-

Figura 13 - Quatro estágios para a extração de fragmentos de curva.



Legenda: Quatro estágios para a extração de fragmentos de curva.
Fonte: O autor, 2017.

mentos de curva provenientes do quarto estágio, embora superiores a outras abordagens, em alguns casos ainda apresentam falta de totalidade, distinção e objetividade. Por isso, características geométricas, fotométricas e uma nova característica de densidade espacial de bordas (*edge sparsity*) foi introduzida em Guo et al. (2014b), Guo, Narayanan e Kimia (2017), sendo combinadas no aprendizado para junção e quebra de curvas (totalidade e distinção) no quinto estágio. Finalmente, a classificação dos fragmentos de curvas entre objetivas e espúrias, removendo falso-positivos é feita no sexto estágio.

O resultado final é um conjunto de curvas altamente representativas e precisas. Estes dois últimos estágios são explorados em profundidade neste trabalho na extração de curvas em imagens de ondas aquáticas. Um importante ponto nessa abordagem multi-estágio é que lacunas maiores do que as presentes em vizinhanças locais estão fora do escopo deste trabalho. No entanto, garantindo-se a qualidade das curvas extraídas, tais lacunas podem ser cobertas em estágios de pós-processamento como o apresentado em Ren, Fowlkes e Malik (2008), Narayanan e Kimia (2012), Usumezbas, Fabbri e Kimia (2017).

No aprendizado com imagens genéricas de Guo et al. (2014b), Guo, Narayanan e Kimia (2017), os algoritmos foram treinados para maximizar o desempenho em:

- correspondência com anotações humanas genéricas em *benchmarks*, Martin, Fowlkes e Malik (2004), Guo e Kimia (2012);
- estabilidade dos fragmentos de curvas detectados entre variações de vistas e iluminação, útil para 3D e aplicações em vídeos.

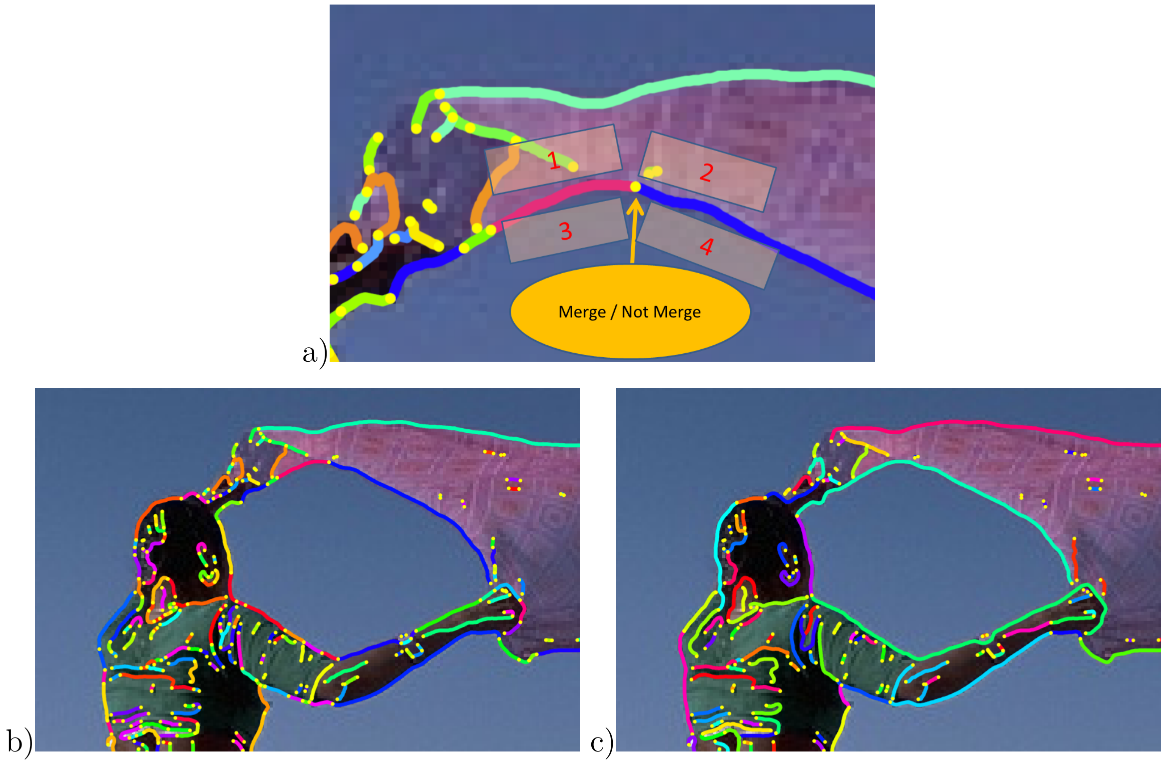
Para casos em geral, esta abordagem de treinamento genérica tem desempenho superior ao estado da arte nas duas formas de avaliação propostas. Esses objetivos não são os mesmos almejados na detecção de ondas aquáticas. No caso tratado aqui, deseja-se uma correspondência específica em uma base de dados feita para este trabalho, detalhada no Capítulo 3. Além disso, podem existir condições particulares desejadas que não se baseiam em aprendizado genérico, mas em considerações referentes à aplicação em si.

Os próximos dois estágios são aqui descritos de forma genérica, salvo algumas poucas considerações.

1.4.1 Estágio V: aprendizado aplicado à junção de fragmentos de curva

O resultado dos estágios anteriores, não-supervisionados, é um conjunto de fragmentos de curvas representados em um grafo, onde curvas são os vértices e junções nas suas extremidades são as arestas. Ainda assim, as curvas extraídas apresentam fragmentações, veja Figura 14. Seguindo a progressão de informação local a global utilizada nos

Figura 14 - A aplicação de um classificador nas possíveis conexões entre as curvas.



Legenda: A aplicação de um classificador nas possíveis conexões entre as curvas, que são representadas por um grafo de conexões, ajuda na redução da fragmentação.
(a) O problema fundamental, conectar ou não duas curvas é representado com as áreas em seu entorno onde as características são extraídas.
(b) Curvas antes da classificação.
(c) Curvas após a classificação.

Fonte: (GUO et al., 2014c).

estágios anteriores baseados em consistência geométrica, este novo estágio utiliza-se de considerações mais globais e não apenas geométricas.

Embora os estágios anteriores às vezes não sejam suficientes para remover todas as ambiguidades, eles produzem resultados que em geral são curvas suficientemente longas para que considerações mais globais e semânticas sejam utilizadas no agrupamento. Neste estágio, procura-se a solução de ambiguidades quanto à junção de contornos: conectá-los ou não. Isto reflete essencialmente um problema de classificação binária, ilustrado na Figura 14 (a).

Um conjunto de oito características fotométricas e geométricas são utilizadas nesta classificação. A relevância destas no ambiente específico de ondas aquáticas é discutida no Capítulo 4.

Características fotométricas

As características fotométricas 1D na imagem ao longo do fragmento de curva extraído no estágio IV, são integrais de funções parametrizadas pelo comprimento de arco desses fragmentos de curva:

- gradiente do brilho — $\int \nabla V(s)\, ds$;
- gradiente da saturação — $\int \nabla S(s)\, ds$;
- gradiente da coloração — $\int \nabla H(s)\, ds$,

onde H, S e V são as componentes da imagem no espaço de cores HSV, descrito em mais detalhadamente na subseção 2.3.2. Essas três são computadas para as duas curvas e comparadas. Uma característica fotométrica 2D analisa a textura das curvas em torno do ponto de encontro:

- diferença de textura — $\chi^2(U_{C^1}, U_{C^2}) + \chi^2(L_{C^1}, L_{C^2})$,

com U_{C^i} e L_{C^i}, respectivamente, o histograma de textura acima e abaixo da curva C^i e $\chi^2(A, B)$ a distância qui-quadrado entre os histogramas A e B dada por

$$\chi^2(A, B) = \sum_{i=1}^{n} \frac{(a_i - b_i)^2}{a_i + b_i}$$

As texturas de diferentes lados da curva são comparadas. Por exemplo, na Figura 14(a), a diferença é feita entre os retângulos 1 e 2, e entre 3 e 4. Essas diferenças se referem à distância qui-quadrado entre os histogramas das texturas em cada região.

Características geométricas

As características geométricas intrínsecas 1D, ao longo da curva, são as tradicionais medidas de:

- curvatura — integral da curvatura absoluta parametrizada pelo comprimento de arco: $\int |\kappa(s)|\, ds$;
- ondulação — a quantidade de pontos de inflexão.

Essas duas características também são calculadas e comparadas entre as duas curvas. A característica medida no ponto de encontro dessas curvas é também chamada de 0D:

- continuidade geométrica — ângulo formado pelas duas curvas no ponto de encontro entre elas.

Finalmente, uma nova característica 2D, introduzida por Guo et al. (2014a), Guo, Narayanan e Kimia (2017), é apresentada:

- esparsidade lateral de *edgels*: definida como a quantidade de *edgels* localizados no entorno de uma curva, normalizada por unidade de comprimento. Estes, quando em baixo número, representam pouca ambiguidade no entorno da curva, assim admite-se que são inversamente proporcionais à chance da curva ser verídica.

1.4.2 Estágio VI: extração de curvas baseada em aprendizagem

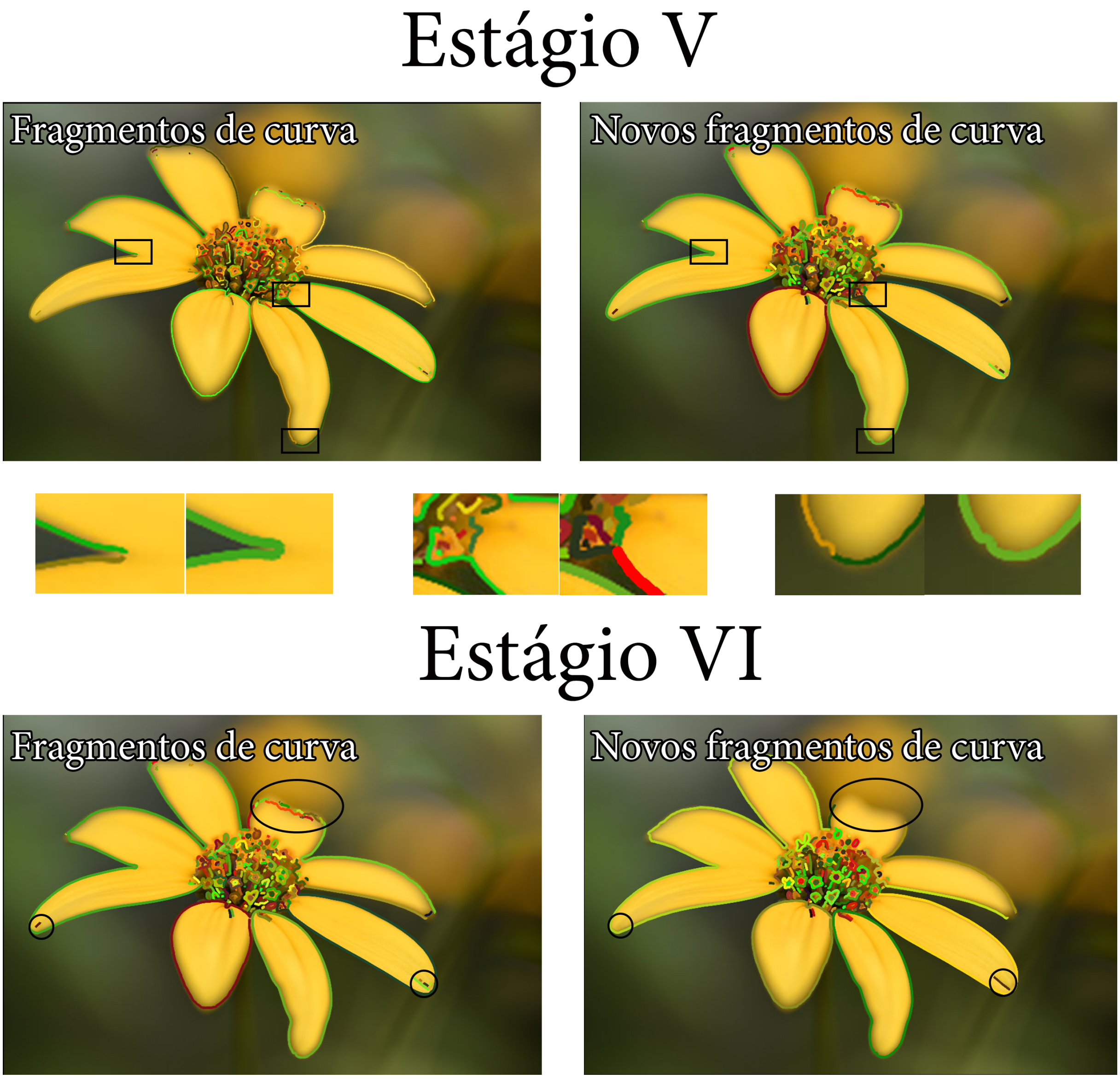
As curvas obtidas pelo método proposto neste trabalho são resultantes da detecção de bordas utilizando limiares de intensidade muito baixos em todo o processo. Praticamente todas as curvas de interesse são capturadas, mas também muitas curvas espúrias, causadas por baixa precisão nas informações e ruídos. Como uma grande quantidade de falso-positivos pode confundir estágios futuros de reconstrução 3D e reconhecimento, este estágio tem por objetivo reduzir os falso-positivos, necessariamente reduzindo também o número de curvas verídicas, isto é, criando falso-negativos.

A redução de falso-positivos é alcançada através do aprendizado sobre curvas verídicas num conjunto de dados de treinamento. Especificamente, este aprendizado ocorre sobre as mesmas características do estágio anterior, excluindo-se as que necessitam de duas curvas para serem calculadas. As características substituídas são a diferença de textura e ângulo entre as curvas, as quais são supridas adicionando-se o tamanho e módulo do gradiente da intensidade de curva detectada. Assim, cada curva é representada como um ponto num espaço de características, sendo este de dimensão elevada.

No conjunto de treino, curvas verídicas e não-verídicas de uma imagem, adquiridas nos estágios anteriores, são divididas em casos positivos e negativos, respectivamente, suas características são computadas e o classificador treinado. A separação destas curvas é feita através da correlação entre as curvas computadas e as curvas de referência (*ground-truth*), através do método proposto por Guo e Kimia (2012).

Para combinar estas características é utilizado o classificador logístico, que atribui uma probabilidade a um contorno de ser verídico ou não. Este tipo de classificador é de fácil treinamento e pouco custoso computacionalmente, sendo excelente para aplicações com alta quantidade de dados, como vídeos. Além disso, a possibilidade de escolha de várias faixas de corte na probabilidade de classificação, permite a análise das curvas de precisão e cobertura que melhor atendem a aplicações específicas.

Figura 15 - Os dois estágios que utilizam o aprendizado de máquina.



Legenda: Os dois estágios que utilizam o aprendizado de máquina.

Fonte: O autor, 2017.