

Homomorfismo entre Grafos para Segmentação e Reconhecimento de Características Faciais em Sequências de Vídeo

MAC 5701 - Tópicos em Ciência da Computação
Monografia Final
Professor Responsável: Yoshiharu Kohayakawa
Departamento de Ciência da Computação
IME-USP

Aluna: Ana Beatriz Vicentim Graciano
abvg@ime.usp.br
Orientador: Roberto Marcondes Cesar Junior
cesar@ime.usp.br

Resumo

Este documento visa apresentar um estudo geral sobre tópicos relativos ao problema de segmentação e reconhecimento de características faciais em sequências de vídeo, solucionado através de uma metodologia baseada em homomorfismo entre grafos, tema do projeto de mestrado da aluna. Em particular, é introduzida a motivação para a análise de tal problema, além de diversas considerações sobre as etapas da metodologia, que constituíram o objeto de estudo da disciplina de *Tópicos em Ciência da Computação*.

Sumário

1	Introdução	3
2	Segmentação e Reconhecimento de Faces através de Homomorfismos entre Grafos	4
2.1	Representação da Face	4
2.2	O Processo de Reconhecimento de Características Faciais . .	6
2.3	O Processo de Rastreamento	7
2.4	Resultados Preliminares	9
3	Segmentação de Imagens: Abordagem através da técnica de Watershed	10
4	Grafos e Morfologia Matemática	12
4.1	Grafos: Conceitos, Definições e Propriedades	12
4.2	Tipos de Grafos e Implementação	13
5	Algoritmos em Grafos	15
6	Conclusão	18

1 Introdução

Faces são elementos muito especiais, pois apresentam papéis importantes em vários aspectos, tais como o social, biológico e teórico. Do ponto de vista cognitivo, a relevância de se estudar faces e as implicações que essas induzem na mente se deve à observação de diversos fatores, dentre os quais os de maior destaque são citados abaixo [8]:

- reconhecimento de faces é uma habilidade adquirida naturalmente pelo ser humano, sendo uma função altamente especializada e que apresenta uma sensível percepção de detalhes e especificidades a partir de estímulos semelhantes (faces e seus componentes não possuem uma grande variabilidade).
- as noções de forma, função e identidade de faces não apresentam uma correlação tão grande quanto em outros tipos de objetos visuais.
- faces transmitem diversos sinais sociais, tais como comportamentais (por exemplo, estado emocional) e fenotípicos (idade, sexo). Portanto, o ser humano pode extrair diferentes tipos de informação a partir de uma face, além da simples identificação desse elemento como tal.

Assim, o estudo e a compreensão dos mecanismos de percepção de faces permitiriam esclarecer questões intrinsecamente ligadas à evolução do cérebro em relação à sua especialização na decodificação de informações carregadas pelas faces, e ao impacto desses sinais sobre nossas crenças e visões acerca do mundo.

No entanto, o estudo de faces não é apenas relevante para as ciências cognitivas, mas também para o âmbito de Visão Computacional, surgindo em diferentes aplicações. Em particular, o tema pode ser decomposto em outros subproblemas: detecção de faces (localização), reconhecimento de faces (identificação ou atribuição de uma face a uma certa pessoa), análise facial (localização de características faciais como íris, narinas, lábios; análise de expressões faciais) e rastreamento de faces e/ou características faciais em seqüências de vídeo.

Em termos de aplicações, grande parte dos esforços concentram-se em trabalhos de interfaces homem-máquina inteligentes (*affective computing*), biometria, reconhecimento e síntese de voz audio-visual, visão robótica, animação de faces mais realistas, teleconferências, análise de bancos de dados multimídia e bases de imagens, sistemas digitais de identificação de pessoas, segurança e vigilância automatizadas e codificação de vídeo digital.

Embora esse t3pico de pesquisa seja de grande interesse na comunidade e apresente diversas produ33es, muitas quest3es permanecem abertas. No caso de sequ4ncias de v3deo, em particular, estimativa de movimento para rastreamento da face (*pose estimation*) e tratamentos para oclus3o parcial ou total de objetos de interesse ainda representam problemas sem solu33es estabelecidas.

Neste documento, o foco principal consiste no problema de reconhecimento de caracter3sticas faciais em sequ4ncias de v3deo. 3 importante ressaltar que o termo reconhecimento 3 utilizado aqui no sentido de classificar uma caracter3stica facial como sendo uma certa componente da face (ex.: l3bios, narinas, sobrancelhas).

Tal problema 3 abordado atrav3s de uma metodologia baseada em homomorfismo entre grafos, sendo uma extens3o da metodologia apresentada para imagens est3ticas em [1, 2].

De modo geral, caracter3sticas faciais s3o modeladas atrav3s de um *grafo relacional de atributos*, ou ARG (do ingl3s, attributed relational graph). Esses grafos s3o obtidos a partir da imagem do gradiente supersegmentado derivado de cada quadro da sequ4ncia. J3 o processo de reconhecimento das caracter3sticas faciais consiste em buscar um homomorfismo apropriado entre o grafo de um quadro da sequ4ncia e aquele de uma imagem modelo em que as caracter3sticas faciais de interesse foram previamente escolhidas.

Uma explica3o detalhada da metodologia 3 apresentada na Se3o 2. Os temas relacionados a esse projeto e que motivaram a pesquisa realizada para a disciplina de t3picos s3o introduzidos nas se33es subsequentes da seguinte forma: Segmenta3o de Imagens: Abordagem atrav3s da t3cnica de Watershed (Se3o 3), Grafos e Morfologia Matem3tica (Se3o 4) e Algoritmos em Grafos (Se3o 5). Finalmente, na Se3o 6, s3o feitas algumas considera33es finais.

2 Segmenta3o e Reconhecimento de Faces atrav3s de Homomorfismos entre Grafos

2.1 Representa3o da Face

Grafos Relacionais de Atributos. Nesta se3o, um grafo dirigido ser3 representado por $\tilde{G} = (N, E)$, onde N representa o conjunto de v3rtices de \tilde{G} e $E \subseteq N \times N$, seu conjunto de arestas. Dois v3rtices a, b de N s3o ditos adjacentes se $(a, b) \in E$. Se todos os v3rtices de \tilde{G} s3o adjacentes entre si, ent3o \tilde{G} 3 denominado completo. Ainda, $|N|$ e $|E|$ denotam, respectivamente, o

número de vértices e de arestas de G .

Um *grafo relacional de atributos* é um grafo a cujos vértices e arestas são associados vetores de atributos. Formalmente, um ARG é definido como $G = (N, E, \mu, \nu)$, onde N e $E \subseteq N \times N$ representam, respectivamente, o conjunto de vértices e de arestas de G . Ainda, $\mu : N \rightarrow L_N$, também chamado de atributo de objeto, é associado a cada vértice de G , enquanto $\nu : E \rightarrow L_E$, ou atributo relacional, é associado a cada aresta.

Dentro desse modelo, vértices representam regiões conexas de características faciais e arestas denotam relações entre as mesmas, ou seja, carregam informações sobre o arranjo espacial de tais características.

Atributos. Para a discussão a seguir, consideremos um ARG $G = (N, E, \mu, \nu)$ e dois vértices quaisquer a, b de N .

O **atributo de objeto** $\mu(a)$ é definido como:

$$\mu(a) = (g(a), w(a), l(a)) \quad (1)$$

O termo $g(a)$ indica o nível de cinza médio da região da imagem associada ao vértice a , enquanto $w(a)$ é um coeficiente obtido através da aplicação de uma wavelet de Morlet. Além disso, $g(a)$ e $w(a)$ são normalizados entre 0 e 1 em relação ao maior nível de cinza possível. Finalmente, $l(a)$ é um rótulo de região.

O **atributo relacional** $\nu(a, b)$, para a, b em E , é definido como:

$$\nu(a, b) = (\vec{v}, sym(a, b)). \quad (2)$$

O atributo \vec{v} denota o vetor $(p_b - p_a)/2d_{max}$, em que d_{max} é a distância máxima entre quaisquer dois vértices do grafo da imagem de entrada enquanto p_a e p_b são os centróides das regiões às quais os vértices a e b correspondem. O termo $sym(a, b)$ representa uma simetria reflexiva calculada como descrito em [3].

O Modelo de Referência da Face. Uma imagem-modelo da face é usada como referência para reconhecer características faciais de interesse. Tipicamente, é utilizado o primeiro quadro da sequência a ser processada, que é segmentado manualmente em regiões relevantes. Ainda, seu respectivo ARG é obtido e deve conter vértices associados a cada característica facial que se deseja rastrear (ex: íris, lábios, sobrancelhas).

2.2 O Processo de Reconhecimento de Características Faciais

Homomorfismo entre Grafos. Considere dois ARGs

$G_1 = (N_1, E_1, \mu_1, \nu_1)$, proveniente de um quadro da sequência, e $G_2 = (N_2, E_2, \mu_2, \nu_2)$, derivado do modelo. Eles serão chamados, respectivamente, de grafo de entrada e grafo do modelo, e terão seus vértices, arcos e atributos indexados pelos mesmos valores que os identificam.

Um *grafo de associação* \tilde{G}_A entre G_1 e G_2 é definido como o grafo completo $\tilde{G}_A = (N_A, E_A)$, onde $N_A = N_1 \times N_2$ e $E_A = E_1 \times E_2$.

Um *homomorfismo* h entre G_1 e G_2 é um mapeamento $h: N_1 \rightarrow N_2$ tal que, $\forall a_1 \in N_1, \forall b_1 \in N_1$, se $(a_1, b_1) \in E_1 \Rightarrow (h(a_1), h(b_1)) \in E_2$. Essa definição pressupõe que todos os vértices em G_1 serão mapeados em G_2 . No caso deste trabalho, tal mapeamento corresponde à fusão de subregiões coerentes na imagem de entrada supersegmentada e, portanto, a obtenção de um homomorfismo apropriado entre G_1 e G_2 é essencial.

Como proposto em [1], uma solução para encontrar um homomorfismo entre G_1 e G_2 pode ser definida como um sub-grafo completo $\tilde{G}_S = (N_S, E_S)$ do grafo de associação \tilde{G}_A , em que $N_S = \{(a_1, a_2), a_1 \in N_1, a_2 \in N_2\}$ é tal que $\forall a_1 \in N_1, \exists a_2 \in N_2, (a_1, a_2) \in N_S$, e $\forall (a_1, a_2) \in N_S, \forall (a_1', a_2') \in N_S, a_1 = a_1' \Rightarrow a_2 = a_2'$, garantindo que cada vértice do grafo de entrada corresponda a exatamente um vértice do grafo do modelo e $|N_S| = |N_1|$. Note que tal solução considera apenas as estruturas de G_1 e G_2 e que há outros homomorfismos possíveis.

Função Critério. Para avaliar a qualidade e a adequação de um dado homomorfismo entre os grafos de entrada e do modelo, é necessário definir uma função critério que considere não somente as estruturas dos grafos, mas também as relações entre os atributos das características faciais. Neste trabalho, a adequação de um homomorfismo é analisada minimizando-se a seguinte função critério:

$$f(\tilde{G}_S) = \frac{\alpha}{|N_S|} \sum_{(a_1, a_2) \in N_S} c_N(a_1, a_2) + \frac{(1 - \alpha)}{|E_S|} \sum_{e \in E_S} c_E(e) \quad (3)$$

onde c_N e c_E são as seguintes medidas de dissimilaridade:

$$c_N(a_1, a_2) = \left\{ \begin{array}{l} \gamma_N |g_1(a_1) - g_2(a_2)| + \\ (1 - \gamma_N) |w_1(a_1) - w_2(a_2)|, \\ \text{se } l(a_1) = l(a_2) \\ \infty, \text{ c.c} \end{array} \right\} \quad (4)$$

$$c_E(e) = \gamma_E \phi_v + (1 - \gamma_E) \phi_{sym} \quad (5)$$

e ϕ_v , ϕ_{sym} são definidos como:

$$\phi_v = \gamma_v \left(\|\vec{v}_1\| - \|\vec{v}_2\| \right) + (1 - \gamma_v) \frac{|\cos \theta - 1|}{2} \quad (6)$$

$$\phi_{sym} = |sym(a_1, b_1) - sym(a_2, b_2)| .$$

Nesse caso, $\cos \theta = \frac{\vec{v}_1 \cdot \vec{v}_2}{\|\vec{v}_1\| \|\vec{v}_2\|}$ e os valores γ_N , γ_E e γ_v são pesos.

Encontrando um Homomorfismo. A procura por um homomorfismo adequado entre os grafos de entrada e do modelo é realizada a partir de um algoritmo de otimização baseado em busca em árvores. Outras alternativas possíveis são abordadas em [1], [4] e [5].

Em linhas gerais, tal algoritmo gera uma árvore de busca em que cada vértice representa um par de vértices (k, l) , $k \in N_1$ e $l \in N_2$. O vértice raiz é rotulado com $(0, 0)$ e expandido em $|N_2|$ filhos rotulados com $(1, l)$, $l = 1 \dots |N_2|$. A cada passo k do algoritmo, o filho que minimiza a função critério, digamos (k, l_{min}) , é escolhido e também expandido em $|N_2|$ filhos $(k + 1, i)$, $i = 1 \dots |N_2|$. Esse processo é repetido até que um vértice $(|N_1|, l)$ seja alcançado, o que garante que todos os vértices de G_1 tenham sido mapeados em um vértice de G_2 , estabelecendo um homomorfismo entre os grafos de entrada e do modelo.

2.3 O Processo de Rastreamento

Nesta seção, serão apresentados o processo geral para rastreamento de características faciais ao longo de uma sequência de vídeo, bem como as contribuições deste trabalho para a generalização do que foi proposto em [1].

Processo Geral. A Figura 1 apresenta a sequência geral de etapas para a realização da segmentação e do reconhecimento de características faciais num quadro qualquer da sequência de vídeo.

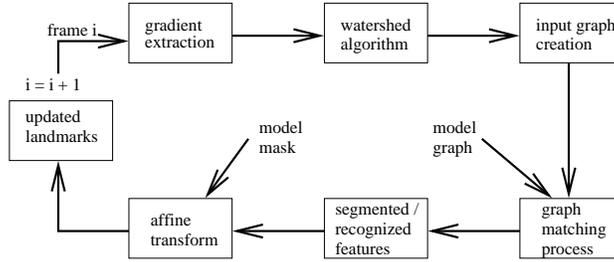


Figura 1: Esquema geral do processo de rastreamento.

Inicialmente, são localizados marcadores aproximados (*landmarks*) para as características faciais de interesse no quadro, os quais serão utilizados posteriormente para restringir a região em que a imagem deverá ser super-segmentada. O passo seguinte consiste na aplicação do algoritmo proposto para imagens estáticas.

Como uma sequência de vídeo possui um caráter dinâmico e suas imagens, em geral, variam em função do tempo, a atualização dos marcadores mostra-se importante para a manutenção da qualidade da segmentação e para o rastreamento adequado das características faciais.

Atualização dos Marcadores. Para aproveitar as informações obtidas diretamente dos grafos e da imagem-modelo na atualização dos marcadores, uma *transformação afim* é aplicada para mapear a imagem-modelo ao quadro atual baseando-se nas características faciais reconhecidas e, então, atualizando-se os marcadores.

No primeiro quadro da sequência, os marcadores previamente obtidos relativos ao modelo são também usados para essa imagem. Nos quadros subsequentes, após o término do processo de reconhecimento, são calculados os centróides das regiões das características faciais de interesse. Além disso, são calculados os centróides das regiões pré-definidas do modelo. Então, a transformação afim que melhor mapeia o conjunto de centróides do modelo para o conjunto relativo ao quadro considerado é estimada e aplicada usando-se a seguinte fórmula [7]:

$$\vec{q} = \alpha(A\vec{s} + \vec{b}) \quad (7)$$

onde A corresponde a uma matriz 2×2 não-singular representando a transformação procurada, α é um valor escalar qualquer, e \vec{q} , \vec{s} são os vetores das coordenadas dos centróides do quadro e do modelo respectivamente.

Essa transformação afim permite, além da atualização dos marcadores, a projeção da imagem-modelo sobre o quadro segmentado e reconhecido, permitindo uma avaliação visual do processo de reconhecimento.

Possíveis Extensões. Embora a atualização dos marcadores permita com que a metodologia reflita no modelo a variação espacial entre os quadros ao longo da sequência de vídeo, nossa pesquisa em andamento objetiva um maior aproveitamento das informações possivelmente redundantes presentes em quadros distintos. Para tanto, foram propostas algumas modificações não somente na estrutura dos ARGs, mas também na forma de se buscar um homomorfismo entre os grafos.

Um ARG poderia ser redefinido como $G = (N, E, \mu, \nu, \delta)$, onde $E = E_{space} \cup E_{time}$ e δ é um vetor de atributos temporais. E_{space} é o conjunto de arestas definido anteriormente, enquanto E_{time} seria o conjunto de *arestas temporais* do grafo. Tais arestas serviriam como medida para a distância entre vértices relacionados provenientes de quadros diferentes da sequência.

A utilidade das arestas temporais fica evidente ao introduzirmos a idéia de grafo *inter-quadros* G^{inter} . Esse grafo seria formado por uma sequência de grafos *intra-quadros* $G_{i-j}^{intra}, \dots, G_i^{intra}, \dots, G_{i+j}^{intra}, j \geq 0$, que constituem ARGs derivados de cada quadro da sequência e que são similares àqueles já utilizados na metodologia, exceto pela existência desse novo conjunto de arestas. Ambos os grafos inter e intra quadros poderiam não ser completos, assumindo uma estrutura mais simples, porém ainda significativa para a extração de relações entre os vértices.

O reconhecimento das características faciais seria feito através da busca por um homomorfismo entre o grafo inter-quadros e o grafo modelo, levando-se em consideração o conjunto de informações carregadas por quadros dentro de uma dada vizinhança e ponderando possíveis influências através do atributo temporal presente na estrutura. Para isso, modificações pertinentes devem ser introduzidas na função critério que buscamos minimizar.

2.4 Resultados Preliminares

Nesta subseção, mostramos alguns dos resultados obtidos a partir da aplicação da metodologia descrita na seção anterior.

Os testes foram realizados a partir de sequências de vídeo com mudanças consideráveis na face ao longo do tempo, tais como sorrisos, piscar de olhos, movimentação da cabeça, entre outras.

A Figura 2 exibe os resultados obtidos da projeção da máscara do modelo sobre alguns quadros de uma sequência composta por 96 imagens coloridas

de tamanho 512×512 , que foram convertidas para níveis de cinza para o propósito do algoritmo.

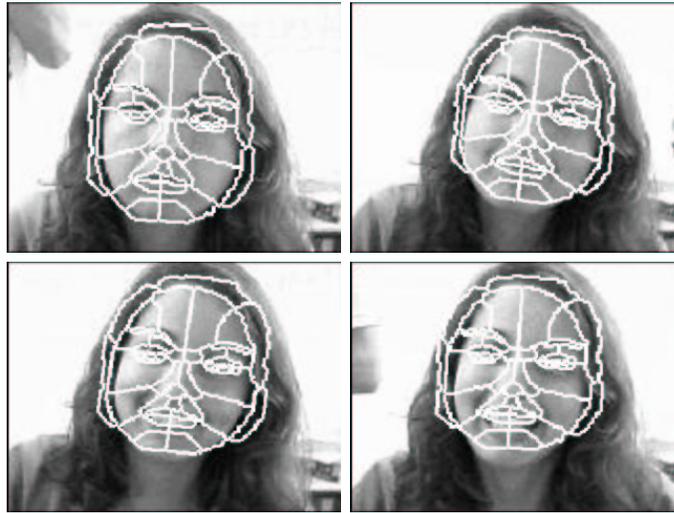


Figura 2: Máscaras do modelo projetadas sobre seus respectivos quadros segmentados e reconhecidos, de acordo com as transformações afins encontradas.

É possível verificar que a máscara do modelo é satisfatoriamente projetada sobre a face, permitindo o rastreamento da mesma ao longo da sequência. Ainda, grande parte das características faciais que podem ser de interesse (ex: sobrancelhas, narinas, nariz, lábios) são corretamente rastreadas, comprovando a validade do processo de reconhecimento e da atualização dos marcadores.

No entanto, certos refinamentos ainda são necessários à metodologia, especialmente em relação a mudanças bruscas entre quadros e às ações que devem ser tomadas quando novas características faciais são introduzidas na sequência.

3 Segmentação de Imagens: Abordagem através da técnica de Watershed

Um dos passos importantes da metodologia apresentada consiste na segmentação dos quadros da sequência através do algoritmo de watershed. A

partir do resultado desse processo de segmentação, extraem-se regiões da imagem que originarão os ARGs de entrada.

O algoritmo de watershed é uma das ferramentas morfológicas mais utilizadas em problemas de segmentação. No entanto, essa técnica é muito sensível a ruídos na imagem, o que resulta numa *supersegmentação* da mesma. No caso do problema de segmentação de características faciais, o número de regiões geradas pelo watershed está diretamente ligado ao número de vértices do grafo relacional de atributos. Portanto, se fosse possível atenuar a supersegmentação, a complexidade do grafo poderia ser reduzida.

O método de watershed foi proposto por Beucher e Meyer em [13] e os autores abordam o problema da supersegmentação de duas formas distintas: através do uso de marcadores para os objetos de interesse, ou através de uma segmentação hierárquica da imagem.

Watershed obtido através de marcadores. A solução por marcadores consiste na escolha de um subconjunto de pixels da imagem relativos aos objetos de interesse. O processo de segmentação através do algoritmo de watershed é então aplicado baseando-se nesse conjunto, restringindo as regiões que permanecerão no resultado. Embora essa abordagem seja eficiente em termos qualitativos, nem sempre é fácil selecionar e extrair marcadores para diferentes imagens. Em outros casos, o problema é a estrutura dos objetos a serem segmentados, os quais podem assumir formas complexas, tamanhos variados, níveis de cinza não uniformes. No caso de sequências de vídeo, seria necessário estimá-los em cada quadro [18], levando em consideração as transições devido ao movimento, mudança de iluminação, entre outros.

Segmentação Hierárquica. A segunda solução proposta pelos autores não exige a existência de marcadores, sendo mais genérica do que a anterior. A idéia central é criar uma *segmentação hierárquica* da imagem, baseada em relações de vizinhança entre as diferentes regiões da imagem.

Para tanto, considere uma imagem f em níveis de cinza e seu gradiente morfológico $g(f)$. O algoritmo de watershed tradicional deve ser aplicado sobre $g(f)$, resultando numa imagem m chamada imagem-mosaico. Cada região de m é rotulada com o valor do nível de cinza de f correspondente ao mínimo de $g(f)$.

A imagem-mosaico contém regiões praticamente homogêneas visualmente, mas que são formadas por pedaços homogêneos da imagem cujos níveis de cinza são muito próximos. Uma região homogênea pode então ser caracterizada por apresentar transições mínimas dentro de sua área. Através dessa

observação, podemos definir o processo de segmentação hierárquica como a fusão de regiões adjacentes de m , iniciando-se o processo a partir daquelas que apresentam transições mínimas.

Para que essa variação do algoritmo de watershed seja implementada, é preciso construir um grafo valorado a partir de m . Os vértices de tal grafo representam conexões entre as regiões adjacentes a uma mesma região original de m , enquanto as arestas carregam informação sobre o módulo da diferença entre os níveis de cinza de tais regiões. Assim, as arestas de menor valor representam as transições mínimas entre diferentes regiões e, a partir delas e dos vértices que estas conectam, é feita a fusão das regiões.

Esse processo de segmentação fornece um subconjunto das fronteiras geradas em m , preservando apenas aquelas que não são circundadas por fronteiras de maior contraste.

A idéia de segmentação hierárquica pode atenuar, portanto, o problema de supersegmentação dos quadros de entrada, preservando a noção de vizinhança entre as regiões de m e permitindo que o processo não dependa do uso de marcadores.

4 Grafos e Morfologia Matemática

Diversos problemas práticos de processamento de imagens podem ser modelados em termos de relações de vizinhança entre regiões. No entanto, a representação convencional de uma imagem por uma matriz de pixels pode ser computacionalmente custosa nesses casos. Portanto, uma forma mais significativa e compacta de representar essas relações de adjacência é através de grafos de vizinhança.

Essa idéia é encontrada na aplicação de grafos como estrutura de dados para diferentes operações da morfologia matemática, tais como erosão, dilatação, abertura e fechamento. No entanto, é necessário estabelecer critérios para criar os grafos a partir de imagens de entrada. Os artigos [11] e [12] descrevem alguns grafos que podem ser gerados a partir de uma partição da imagem.

4.1 Grafos: Conceitos, Definições e Propriedades

A definição de grafo utilizada é a usual. Um grafo G é um par (V, E) , onde V é o conjunto de vértices e E é o conjunto de arestas. Em geral, as informações armazenadas nos vértices estão ligadas a propriedades das regiões a que correspondem e podem ser univalueadas ou multivalueadas. Em imagens binárias, os vértices representam tanto objetos (rótulo 1) quanto fundo

(valor 0). No caso de imagens em níveis de cinza, os vértices podem conter o valor de nível de cinza médio das regiões que representam. As arestas do grafo também podem ser valoradas, embora não carreguem informações na discussão presente.

Os grafos considerados para modelar as noções de vizinhança apresentam as seguintes propriedades:

- são não-orientados (devido à simetria das relações de vizinhança)
- são simplesmente conexos, ou seja, existe apenas uma única aresta ligando cada par de regiões
- não possuem laços

O conjunto de vizinhos N_E de um vértice $v_i \in V$ é definido como o conjunto de vértices v_j tal que:

$$N_E(v_i) = \{v_j \in V, (v_i, v_j) \in E\} \quad (8)$$

Um caminho de comprimento k entre quaisquer dois vértices v e $w \in V$ é uma $k + 1$ -tupla (a_0, a_1, \dots, a_k) tal que $v = a_0$, $w = a_k$ e $(a_{i-1}, a_i) \in E \forall i$, $i = 1, \dots, k$.

A distância d entre um par de vértices $(v, w) \in V^2$ é definida como o comprimento do caminho mínimo existente entre ambos. Há também a noção de vizinhos de ordem- k de um vértice $v \in V$, os quais formam o conjunto de vértices w que distam k arestas de v , ou seja, $d(v, w) = k$.

4.2 Tipos de Grafos e Implementação

Para que um grafo seja usado para modelar as relações de vizinhança entre componentes da imagem, é preciso definir uma estrutura para o mesmo e derivar uma implementação correspondente a essa estrutura.

Supondo que os vértices de V representam objetos da imagem, podemos definir alguns grafos válidos, porém não necessariamente conexos ou planares:

- k vizinhos mais próximos: um vértice v é ligado por uma aresta a cada um dos k vértices mais próximos a ele.
- objetos à distância d de um vértice v : uma aresta é criada entre um vértice v e cada um dos vértices que distam dele, no máximo, d unidades. A métrica aqui utilizada é a Euclidiana.

- vizinhos baricêntricos: se v está contido no casco convexo formado por um certo conjunto de vértices, então esses pontos são os vizinhos baricêntricos de v . Esta construção não garante a unicidade do grafo.

Embora os grafos supracitados sejam pertinentes, há um outro conjunto de grafos derivados do conceito de *Diagrama de Voronoi* que se mostram mais robustos e interessantes, pois não dependem de parâmetros como nos casos anteriores, e são planares e conexos.

Antes de introduzi-los, no entanto, é preciso definir o que é o Diagrama de Voronoi de um conjunto V .

Se V é considerado um conjunto finito de pontos no plano, então o diagrama de Voronoi correspondente a ele é dado pela união dos lugares geométricos de influência de cada um dos pontos de V :

$$\bigcup_p \{Z(p) = \{m \in \mathbb{R}^2, \forall q \in V \setminus \{p\}, \text{dist}(m, p) < \text{dist}(m, q)\}\} \quad (9)$$

No caso geral, $V = \{C_1, \dots, C_n\}$, onde C_i é um conjunto disjunto e compacto, e as definições de distância envolvidas são aquelas aplicáveis a conjuntos. Podemos notar que tal diagrama define uma partição sobre o domínio analisado.

Finalmente, os grafos derivados dessa formulação são apresentados abaixo.

Triangulação de Delaunay. Também conhecido como grafo dual do Diagrama de Voronoi, este grafo é constituído por todos os pares de pontos (p, q) cujos polígonos de Voronoi associados ($Z(p)$ e $Z(q)$) são adjacentes. Assim, arestas em G denotam adjacências entre regiões.

Grafo de Gabriel. As arestas deste grafo são geradas se, para um par de pontos $(p, q) \in V^2$, a seguinte relação é satisfeita:

$$\forall m \in V \setminus \{p, q\}, \text{dist}(m, p)^2 + \text{dist}(m, q)^2 > \text{dist}(p, q)^2 \quad (10)$$

Grafo de Vizinhaça Relativa. As arestas geradas para este grafo são tais que, para um par de pontos $(p, q) \in V^2$, vale;

$$\forall m \in V \setminus \{p, q\}, \text{dist}(p, q) < \begin{cases} \text{dist}(m, p) \\ \text{ou} \\ \text{dist}(m, q) \end{cases} \quad (11)$$

Árvore Geradora Mínima. Esse grafo é composto por todos os vértices de V , sendo que sua estrutura é uma árvore de custo/comprimento mínimo. Se as arestas forem valoradas, então tal árvore deverá conter todos os vértices de V e um subconjunto de arestas tais que a soma de seus valores é o mínimo possível.

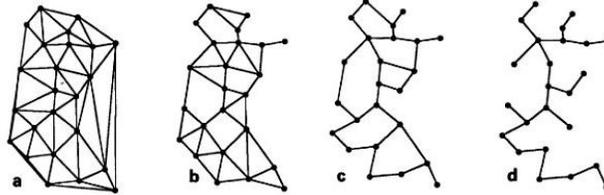


Figura 3: Grafos derivados de diagramas de Voronoi: (a) Triangulação de Delaunay; (b) Grafo de Gabriel; (c) Grafo de Vizinhança Relativa (d) Árvore Geradora Mínima.

A implementação do algoritmo para gerar o diagrama de Voronoi é amplamente encontrada na literatura de Geometria Computacional [10] e possui complexidade $O(n \log n)$. Já a derivação dos grafos a partir do diagrama consome tempo proporcional a $O(n)$, onde n é o número de pixels da imagem.

Tais estruturas seriam úteis para redefinir o tipo de grafo utilizado como ARG. Ao invés de se criar um grafo completo da face, poderia ser tomado um grafo de regiões adjacentes presentes na mesma. Essa estrutura forneceria informações a respeito da relação espacial entre certas regiões e também implicaria em um custo menor de armazenamento dos ARGs, o que é interessante para a versão da metodologia que considera grafos inter-quadros e intra-quadros. Além disso, o resultado do algoritmo de watershed está relacionado ao conceito do diagrama de Voronoi, o que facilitaria a construção de qualquer um dos grafos derivados dessa partição da imagem.

5 Algoritmos em Grafos

A utilização de grafos também é bastante expressiva em outros problemas de Processamento de Imagens, Visão Computacional e Reconhecimento de Padrões.

Algoritmos	Aplicação	Vantagens vs. Desvantagens
Buscas em largura e profundidade	Funções-distância simples [11] Conectividade Cálculo do número de Euler como descritor de formas [9]	Algoritmos rápidos e simples Implementações através de filas ou pilhas Função distância restrita
Dijkstra, Programação Dinâmica	Distâncias euclidianas e geodésicas [19] Segmentação de imagens 2D rastreamento visão estéreo	Relativa simplicidade flexibilidade
Cortes mínimos	Segmentação em quaisquer dimensões agrupamento de características superfícies minimais visão estéreo	Algoritmos razoavelmente complexos algoritmos lentos nicho de pesquisa em expansão
Árvores/Florestas Geradoras Ótimas	Watersheds crescimento de regiões a partir de semente agrupamento de características segmentação	Algoritmos gulosos (<i>Priority First Search</i>) Simples, porém podem não resultar na melhor solução, sendo sensíveis a erros e ruídos
Isomorfismo/Homomorfismo entre grafos	Reconhecimento de formas em geral	Não existem algoritmos polinomiais para a solução de tais problemas (problema intratável)
<i>Matching</i> ótimo e classificações	Rastreamento de múltiplos objetos	Simplificação do Problema de Corte Mínimo

Tabela 1:

Como dito anteriormente, a representação através de grafos é mais compacta em diversas situações e apresenta a vantagem de se reaproveitar algoritmos já conhecidos da Teoria dos Grafos dentro do contexto dessas áreas.

A Tabela 1 introduz alguns algoritmos em grafos e suas respectivas aplicações.

Um estudo aprofundado da relação entre o problema de casamento entre grafos e a área de Reconhecimento de Padrões é feito no artigo [16], bem como em [17]. Os autores discutem a relevância atual do uso de grafos em problemas de classificação e apresentam uma taxonomia para as diversas formas de se realizar o processo de casamento entre grafos, que consiste no estabelecimento de uma correspondência entre vértices e arcos de dois grafos. Um breve resumo das técnicas é mostrado a seguir:

- *isomorfismo entre grafos*: consiste no problema mais simples de casamento entre grafos, pois implica na busca por um mapeamento bijetor entre os nós de ambos os grafos, de forma a preservar as arestas. Esse tipo de mapeamento não é de grande utilidade em reconhecimento de padrões por não ser tolerante a variações estruturais. Como os grafos geralmente representam padrões similares mas resultantes de alguma perturbação, é virtualmente impossível encontrar um isomorfismo entre eles.
Uma variação deste método de mapeamento é o isomorfismo de subgrafos, em que se busca um mapeamento bijetor apenas entre um dos grafos originais e um subgrafo do outro.
- *Máximo Subgrafo Comum*: tal técnica é mais robusta do que as anteriores e consiste na busca pelo maior subgrafo existente em um dos grafos iniciais que seja isomórfico a um subgrafo do outro. Esse problema é equivalente à determinação de um clique máximo num grafo de associação gerado a partir dos grafos iniciais.
- *Distância de edição entre grafos*: para tolerar variações significativas entre os grafos devido a ruídos ou oclusões, foram propostos algoritmos baseados em modelos de erro e medidas de números de edições necessárias para transformar um certo grafo em outro.
- *algoritmos inexatos ou subótimos*: baseiam-se em heurísticas para estimação do resultado global ótimo e rodam em tempo polinomial, ao contrário de grande parte dos algoritmos citados anteriormente, que são exponenciais.

O destaque recebido pelo problema de casamento entre grafos comprova a validade do processo de reconhecimento de características faciais adotados na metodologia proposta.

6 Conclusão

O problema de segmentação e reconhecimento de características faciais pertence ao estado-da-arte do estudo de faces em visão computacional. Além disso, a metodologia escolhida para abordar a questão utiliza-se de técnicas de classificação (busca por homomorfismos entre grafos) e modelagem da face (uso de grafos) que se enquadram numa tendência de estudo atual e em contínua expansão.

Os tópicos estudados também elucidaram questões sobre a relevância do tema de pesquisa e o conhecimento de técnicas alternativas para solucioná-lo. Além disso, permitiram coletar novas idéias em relação às estruturas possíveis dos grafos e aos diversos processos de classificação existentes baseados em algoritmos em grafos.

Referências

- [1] R. Cesar, E. Bengoetxea and I. Bloch. Inexact graph matching using stochastic optimization techniques for facial feature recognition. In *16th International Conference on Pattern Recognition*, volume 2, pages 465-468, August 2002.
- [2] R.M. Cesar Jr. and I. Bloch. First results on facial feature segmentation and recognition using graph homomorphisms. In *Proc. VI Simpósio Ibero-Americano de Reconhecimento de Padrões*, pages 95-99, Florianópolis, Brazil, 2001.
- [3] O. Colliot, A.V. Tuzikov, R.M. Cesar Jr. and I. Bloch. Approximate reflectional symmetries of fuzzy objects with an application in model-based object recognition. In Press. *Fuzzy Sets and Systems*, 2003.
- [4] A. Perchant and I. Bloch. Fuzzy morphisms between graphs. *Fuzzy Sets and Systems*, 128(2):149-168, 2002.
- [5] R.C. Wilson and E.R. Hancock. A Bayesian compatibility model for graph matching. In *Pattern Recognition Letters*, 17(3):263-276, 1996.
- [6] V. Kruger and G. Sommer. Affine real-time face tracking using a wavelet network. In *Proc. of ICVV'99 Workshop Recognition, Analysis, and Tracking of Faces and Gestures in Real-Time Systems*, pages 141-148, Corfu, Greece, 1999.
- [7] R.M. Cesar Jr. and L. da F. Costa. *Shape Analysis and Classification - Theory and Practice*. CRC Press, 1 edition, 2001.
- [8] A. W. Young. *Face and Mind*. Oxford Cognitive Science, Oxford, 1998.
- [9] T.Y. Kong and A. Rosenfeld. Digital Topology: Introduction and Survey. In *Computer Vision, Graphics and Image Processing*, vol. 48, pp. 357-393, 1989.
- [10] J.D. Boissonnat and M. Yvinec. *Géométrie Algorithmique*. Ediscience International, 1995.
- [11] L. Vincent. Graphs and Mathematical Morphology. In *Signal Processing*, vol. 16, pp. 365-368, 1989.

- [12] L. Vincent. *Algorithmes Morphologiques A Base de Files D'Attente et de Lacets. Extension aux Graphes*. These présentée à l'École Nationale Supérieure des Mines de Paris par obtenir le titre de Docteur em Morphologie Mathématique, 1990.
- [13] S. Beucher and F. Meyer. The Morphological Approach to Segmentation: The Watershed Transform. In *Mathematical Morphology in Image Processing*, Marcel Dekker Inc., pp. 433-481, 1993.
- [14] A. Murat Tekalp. *Digital Video Processing*. Prentice Hall Signal Processing Series, Prentice Hall, 1995.
- [15] M. Shah and R. Jain. *Motion-Based Recognition*. Computational Imaging and Vision Series, Kluwer Academic Publishers, 1997.
- [16] M. Vento, D.Conte, et. al. Thirty years of Graph Matching in Pattern Recognition. To appear in *International Journal of Pattern Recognition and Artificial Intellifence*.
- [17] M. Vento, D.Conte, et. al. Graph Matching Applications in Pattern Recognition and Image Processing. In *Proceedings of ICIP'03* .
- [18] F.C. Flores. Segmentação de Seqüências de Imagens por Morfologia Matemática. Dissertação de Mestrado, IME, São Paulo, 2000.
- [19] J.Serra. *Image Analysis and Mathematical Morphology*. Computational Imaging and Vision, Vol. 2, Academic Press, 1997.