

2

Visão Geral das Galerias Inteligentes

Neste capítulo descrevemos as Galerias Inteligentes e enfatizamos sua aplicabilidade em diversos problemas em Computação Visual. Começaremos descrevendo as Galerias de Design, que são interfaces para gerar parâmetros como, por exemplo, posição de câmera. Em seguida, mostramos como as Galerias Inteligentes incorporam aprendizagem nas Galerias de Design, com o objetivo de classificar e selecionar parâmetros de acordo com as preferências subjetivas do usuário. Finalmente, tratamos da aplicabilidade desta nova abordagem, ilustrando como as Galerias Inteligentes podem ser utilizadas em diversos contextos.

2.1

Galerias de Design

As Galerias de Design foram introduzidas para ajudar o usuário em problemas de Computação Visual que envolvem o ajuste de um grande número de parâmetros para se obter um resultado satisfatório. Podemos dar como exemplos posicionamento de câmera e luz, extração de isosuperfícies a partir de dados volumétricos, funções de transferência de cor para renderização de volumes, entre outros. A otimização dos parâmetros desses problemas é muitas vezes subjetiva e trabalhosa. Um problema que tratamos com mais detalhe nesta tese é do problema do fotógrafo, que consiste em escolher a melhor posição de câmera.

Seja $\Omega \subset \mathbb{R}^n$ o espaço dos parâmetros ω (e.g. a posição da câmera), e Π o espaço dos resultados, tipicamente imagens ou representações gráficas. É importante destacar que, em geral, os parâmetros ω são multi-dimensionais, o que complica a busca de soluções. Ademais, a função $\pi: \Omega \mapsto \Pi$ que mapeia parâmetros em resultados é muitas vezes não linear e descontínua, dificultando o trabalho do usuário. Isso é o caso do *rendering*, onde Ω é um espaço que pode incluir, por exemplo, a posição da câmera e sua direção de visão. Além disso, a avaliação de π por si só pode ser computacionalmente cara.

As Galerias de Design (26) foram introduzidas em 1997 com o objetivo

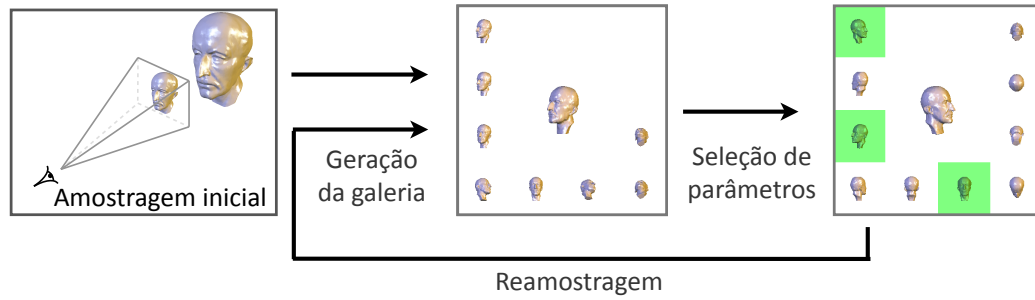


Figura 2.1: Galerias de Design: uma galeria é gerada a partir de uma amostragem $\Psi \subset \Omega$. O usuário seleciona bons parâmetros e opcionalmente uma nova galeria é gerada com parâmetros de Ω que refinam as seleções do usuário.

de dar assistência ao usuário na seleção de parâmetros quando o resultado puder ser representado por meio de imagens, ou seja, quando Π for um espaço de imagens. Nesta metodologia, realiza-se inicialmente uma amostragem do espaço dos parâmetros $\Psi \subset \Omega$. O conjunto de amostras é dado de entrada para a função π , e finalmente o conjunto das imagens resultantes $\pi(\Psi)$ é organizado em uma interface onde o usuário pode interagir selecionando parâmetros. Opcionalmente, uma reamostragem no espaço dos parâmetros é realizada considerando-se a seleção do usuário e seguindo-se alguma estratégia de refinamento. No trabalho original (26), este refinamento é realizado por *clustering*. A figura 2.1 descreve as Galerias de Design.

2.2 Galerias Inteligentes

Podemos observar que os parâmetros selecionados pelo usuário nas suas interações com Galerias de Design formam um rico banco de dados que caracteriza suas preferências. Desse ponto de vista, um método que seja capaz de interpretar esse banco de dados será capaz também de aprender as preferências do usuário, e poderá ajudá-lo a resolver problemas de ajuste de parâmetros que respeitem suas preferências subjetivas.

Com este objetivo, as Galerias Inteligentes estendem o uso das Galerias de Design realizando aprendizagem supervisionada, ou seja, aprendizagem por exemplo, a partir deste banco de dados. Em particular, esta abordagem é apropriada para resolver problemas subjetivos. A figura 2.2 exhibe os principais componentes das Galerias Inteligentes.

Na nova abordagem, os parâmetros são classificados pelo usuário como bons ($c_i = +1$) ou ruins ($c_i = -1$). Esta classificação é processada e incor-

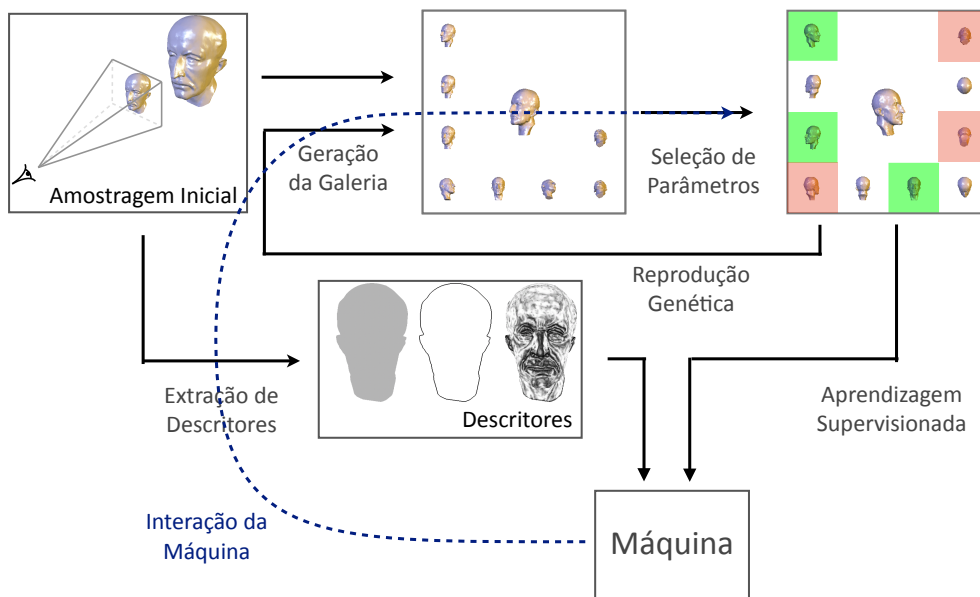


Figura 2.2: Galerias Inteligentes: uma galeria é gerada a partir de uma amostragem $\Psi \subset \Omega$. O usuário classifica imagens $\pi(\omega_i) \in \pi(\Psi)$ como boas (verde) ou ruins (vermelho), ou simplesmente ignora (branco). Os parâmetros selecionados são incorporados a uma máquina de aprendizagem e usados posteriormente para ajudar o usuário a escolher parâmetros.

porada a um banco de dados de treinamento (*training set*). O processamento dos parâmetros classificados pelo usuário envolve a extração de descritores, que representam as preferências do usuário de maneira mais explícita. Observe que o próprio parâmetro ω_i geralmente não pode ser usado como descritor, pois pode não ser intrínseco da instância do problema. A etapa de extração de descritores será explicada detalhadamente no capítulo 5.

A partir das seleções do usuário, uma nova galeria é gerada, refinando a seleção do usuário na galeria anterior. Essa geração pode ser feita por um algoritmo genético que realiza combinações convexas no espaço dos parâmetros (ver apêndice B). Porém, outras estratégias de refinamento poderiam ser utilizadas.

Usando o banco de dados atualizado, a máquina de aprendizagem é capaz de imitar o usuário, ordenando a galeria, colocando bons parâmetros em posições privilegiadas, ou realizando uma seleção automática dos melhores parâmetros. Para esta função, usamos neste trabalho uma adaptação das Máquinas de Suporte Vetorial (45). Justificamos sua escolha pelas suas carac-

terísticas e qualidades que atendem aos requisitos de nossa abordagem. Estas serão explicitadas mais adiante.

Uma introdução às Máquinas de Suporte Vetorial é dada no apêndice A. A adaptação desenvolvida neste trabalho é apresentada no capítulo 3 e analisada no capítulo 4.

2.3

Aplicabilidade

As Galerias Inteligentes se apresentam como uma boa alternativa para resolver problemas subjetivos em Computação Visual, devido a sua capacidade de aprender preferências do usuário por meio de exemplos. Mencionamos aqui alguns destes problemas e desenvolvemos um em particular no final deste trabalho.

Exemplos de problemas nos quais as Galerias Inteligentes se aplicam incluem:

1. Edição de imagens: As Galerias Inteligentes podem ser aplicadas para aprender preferências envolvendo diversos parâmetros de edição de imagens, como brilho, contraste, saturação, parâmetros específicos de filtros, entre outros.
2. Extração de isosuperfícies: Um problema bastante relevante em Computação Gráfica trata de descobrir os melhores isovalores de uma função definida sobre um volume com o objetivo de extrair uma isosuperfície.
3. Iluminação de cenas tridimensionais: Dada uma cena tridimensional, as Galerias Inteligentes podem ser utilizadas para aprender e aplicar parâmetros de luz como posição, direção, intensidade e tipo de luz.
4. Funções de transferência em volumes: As Galerias Inteligentes podem ser utilizadas para, dada uma função de densidade definida sobre um volume, determinar funções de transferência de cor para renderização do volume, provendo um melhor entendimento do objeto gráfico (35, 27, 31).
5. Otimização de posicionamento de câmera (problema do fotógrafo): Dado um modelo tridimensional, podemos usar as Galerias Inteligentes para resolver o problema de escolher parâmetros de câmera que renderizem uma imagem bidimensional ótima. No caso de animações tridimensionais, como em simulação de multidões (30) e jogos (13), este problema torna-se mais complexo e laborioso, pois é necessário escolher parâmetros de câmera em cada quadro da animação, respeitando a coerência temporal.

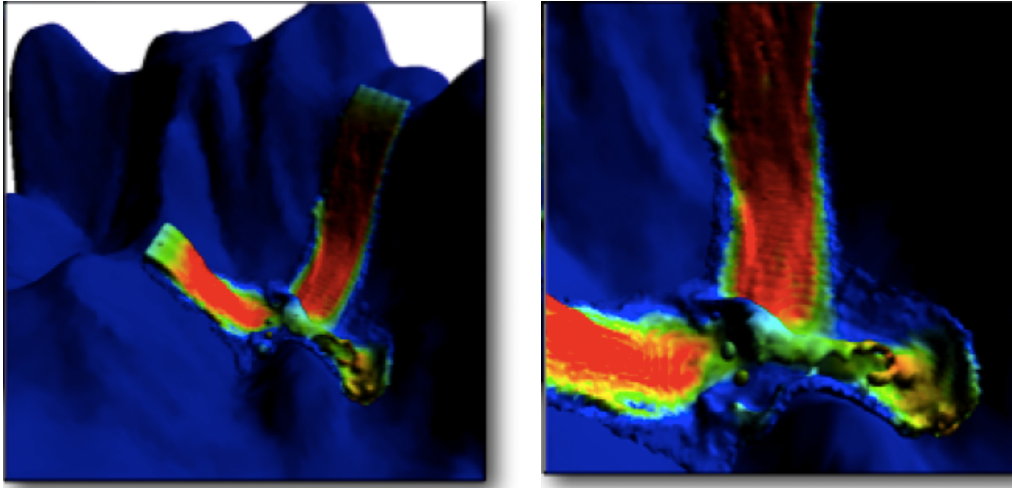


Figura 2.3: Subjetividade: melhor visão de um quadro de uma simulação de fluídos de acordo com as preferências de um designer (esquerda) e de um pesquisador na área (direita).

Trataremos neste trabalho especialmente de otimização de posicionamento de câmera em cenas tridimensionais. Este problema é altamente subjetivo, pois cada usuário possui suas próprias preferências, que podem mudar também de acordo com a aplicação, como ilustrado na figura 2.3. Outra característica importante é sua natureza não-linear: pequenas alterações de parâmetros, como a posição da câmera, podem alterar completamente a avaliação do usuário sobre a visão resultante. Isto justifica o uso de aprendizagem não linear, usando Máquinas de Suporte Vetorial, por exemplo.

Neste problema, o espaço de parâmetros Ω é composto pelo espaço da câmera, o qual é multidimensional, em geral. Dependendo da aplicação, este espaço pode ser composto não somente pela posição da câmera, mas também pela direção de visão, orientação da câmera (*up vector*), entre outros. Em nosso trabalho, dependendo da aplicação, alguns parâmetros de câmera são mantidos fixos para diminuir a dimensão do problema. A função de mapeamento π será simplesmente uma função de *rendering*, que recebe atributos da câmera e renderiza uma imagem contida no espaço de resultados Π . Esta função pode ser extremamente descontínua, principalmente em cenas com grande potencial de oclusão.

Com estas características, o problema de encontrar bons pontos de vista mostra-se apropriado para ser explorado com Galerias Inteligentes. A figura 2.4 exhibe um exemplo de galeria gerada a partir de uma amostragem uniforme em um espaço de câmera. A partir desta galeria inicial, o usuário seleciona os parâmetros, ou pontos de vista, classificando-os como bons ou ruins (verde e vermelho respectivamente na figura 2.4), ou ignorando-os. A partir desta

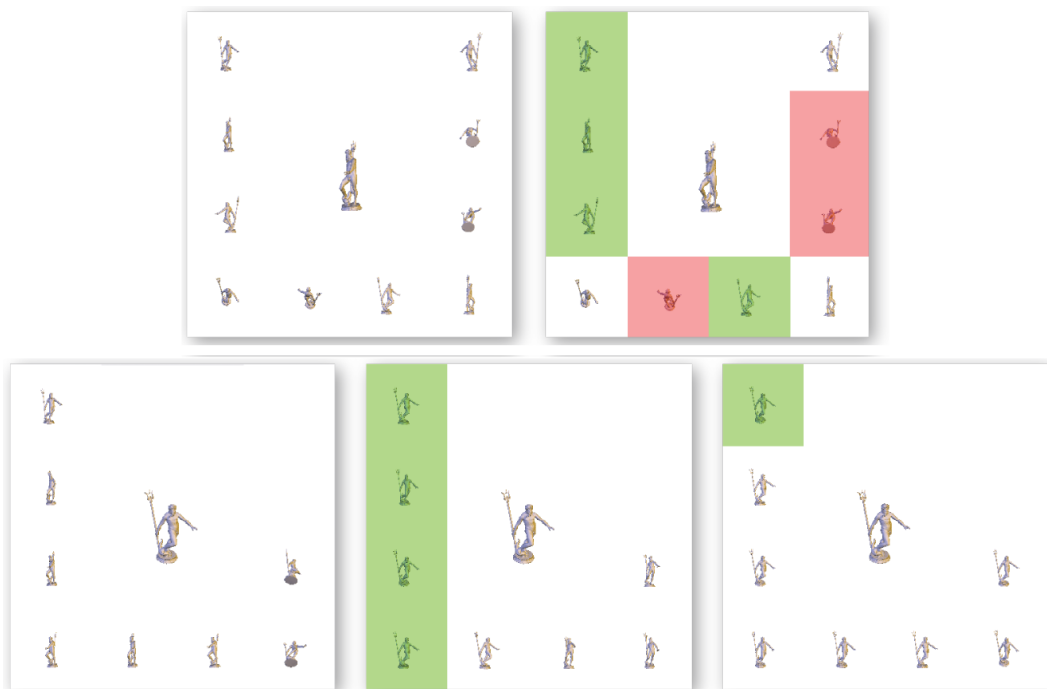


Figura 2.4: Exemplo do uso das Galerias Inteligentes: a partir de uma galeria inicial (superior esquerda), o usuário seleciona e marca parâmetros bons e maus (superior direita); uma nova galeria é gerada (inferior esquerda) e o usuário continua o processo (inferior meio) até chegar a uma visão final satisfatória (inferior direita).

seleção, uma nova galeria é gerada usando um algoritmo genético. A máquina então usa as seleções do usuário para aprender suas preferências, e em seguida ordena e seleciona as melhores visões. Este processo continua até que o usuário encontre uma visão final satisfatória.

Finalmente, o uso das Galerias Inteligentes pode acelerar bastante processos como, por exemplo, determinar as melhores visões em todos os quadros-chaves de uma animação. Neste caso, o usuário poderia selecionar visões de alguns *keyframes* usando as Galerias Inteligentes, que em seguida seriam capazes de selecionar automaticamente as visões dos *keyframes* subsequentes. Experimentos desta natureza são realizados e apresentados no capítulo 6.