



**Thales Miranda de Almeida Vieira**

**Galerias Inteligentes e otimização de  
posicionamento de câmera**

**Tese de Doutorado**

Tese apresentada ao Programa de Pós-graduação em Matemática Aplicada do Departamento de Matemática da PUC-Rio como requisito parcial para obtenção do título de Doutor em Matemática Aplicada

Orientador : Prof. Thomas Lewiner  
Co-Orientador: Prof. Adailson Peixoto  
Co-Orientador: Prof. Luiz Velho

Rio de Janeiro  
Janeiro de 2010



**Thales Miranda de Almeida Vieira**

**Galerias Inteligentes e otimização de posicionamento de câmera**

Tese apresentada ao Programa de Pós-graduação em Matemática Aplicada do Departamento de Matemática do Centro Técnico Científico da PUC-Rio como requisito parcial para obtenção do título de Doutor em Matemática Aplicada. Aprovada pela comissão examinadora abaixo assinada.

**Prof. Thomas Lewiner**

Orientador

Departamento de Matemática — PUC-Rio

**Prof. Adailson Peixoto**

Co-Orientador

Instituto de Matemática – UFAL

**Prof. Luiz Velho**

Co-Orientador

IMPA

**Prof. Claudio Esperança**

COPPE – UFRJ

**Prof. Esteban Clua**

Instituto de Computação – UFF

**Prof. Ricardo Marroquim**

COPPE – UFRJ

**Prof. Geovan Tavares**

Departamento de Matemática – PUC-Rio

**Prof. Sinésio Pesco**

Departamento de Matemática – PUC-Rio

**Prof. José Eugenio Leal**

Coordenador do Centro Técnico Científico — PUC-Rio

Rio de Janeiro, 15 de Janeiro de 2010

Todos os direitos reservados. Proibida a reprodução total ou parcial do trabalho sem autorização da universidade, do autor e do orientador.

### **Thales Miranda de Almeida Vieira**

Graduou-se em Ciência da Computação na Universidade Federal de Alagoas. Concluiu o Mestrado em Matemática na Universidade Federal de Alagoas. Sua dissertação foi intitulada "Registro Automático de Superfícies Usando Spin Images", tendo como área de concentração Computação Gráfica. Desde então, vem trabalhando em diversos temas durante a preparação do doutorado no Laboratório Matmídia da PUC-Rio, incluindo Posicionamento de Câmera, Registro de Superfícies, Parametrização de Superfícies, Deformação de Superfícies e Video 3D, gerando trabalhos publicados. Também participou de dois projetos de pesquisa para a Petrobras: Netgocad e Kernelsis, estando este último diretamente relacionado com o tema desta tese.

#### Ficha Catalográfica

Vieira, Thales

Galerias Inteligentes e otimização de posicionamento de câmera / Thales Miranda de Almeida Vieira; orientador: Thomas Lewiner; co-orientadores: Adailson Peixoto, Luiz Velho. — Rio de Janeiro : PUC-Rio, Departamento de Matemática, 2010.

v., 78 f: il. ; 29,7 cm

1. Tese (Doutorado em Matemática Aplicada) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Matemática.

Inclui referências bibliográficas.

1. Matemática – Tese. 2. Galerias Inteligentes. 3. Galerias de Design. 4. Computação Gráfica. 5. Aprendizagem. 6. Posicionamento de Câmera. 7. Algoritmos Genéticos. 8. Máquinas de Suporte Vetorial (SVM). I. Lewiner, Thomas. II. Peixoto, Adailson. III. Velho, Luiz. IV. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Matemática. V. Título.

CDD: 510

## Agradecimentos

A meus pais, João e Magnólia, por todo tipo de apoio;

A meu orientador, Thomas Lewiner, por todo apoio, dedicação, paciência, incentivo e amizade cultivada nesses anos;

A meus co-orientadores Adailson Peixoto e Luiz Velho, por todo apoio e pelas discussões produtivas;

À Sara, pelo carinho, paciência, incentivo e perseverança nestes anos;

A todos os professores do Laboratório Matmídia, especialmente os professores Hélio Lopes e Geovan Tavares pelas colaborações no desenvolvimento deste trabalho;

A meus colegas Alex Bordignon e Renner Castro pelas colaborações no desenvolvimento deste trabalho;

A todos os meus colegas do Laboratório Matmídia pela amizade e por todo e qualquer tipo de colaboração nesses anos;

A todos os funcionários do Departamento de Matemática da PUC-Rio;

À PUC-Rio, pela bolsa de isenção VRAC;

À CAPES, pela bolsa de doutorado.

## Resumo

Vieira, Thales; Lewiner, Thomas; Peixoto, Adelailson; Velho, Luiz. **Galerias Inteligentes e otimização de posicionamento de câmera**. Rio de Janeiro, 2010. 78p. Tese de Doutorado — Departamento de Matemática, Pontifícia Universidade Católica do Rio de Janeiro.

A Computação Visual levanta vários problemas envolvendo um grande número de parâmetros e tendo uma forte componente subjetiva. Um exemplo de tal problema é o do fotógrafo, tentando otimizar a posição, orientação e abertura da câmera para capturar uma cena tridimensional. As Galerias de Design são comumente usadas em tais problemas para apoiar o usuário na busca da melhor combinação de parâmetros. Porém, a escolha repetida dessas combinações de parâmetros em várias instâncias do problema torna as Galerias de Design usuais laboriosas e não aproveita as interações passadas do usuário para ajudá-lo. Este trabalho apresenta as Galerias Inteligentes, uma nova abordagem para realizar aprendizagem de problemas subjetivos, tais como posicionamento de câmera. A interação do usuário com Galerias de Design ensina uma máquina de aprendizagem estatística. A máquina treinada é capaz de imitar o usuário, classificando parâmetros ou selecionando automaticamente o melhor parâmetro. O processo de aprendizagem baseia-se nas Máquinas de Suporte Vetorial (SVM), que realiza a classificação de parâmetros representados por uma coleção de descritores. Realizamos experiências com o problema de otimização de posicionamento de câmera, que demonstraram que a técnica proposta é eficiente e lida com cenas apresentando elevada oclusão e complexidade de profundidade. Este trabalho também inclui validações de usuário da interface da galeria inteligente.

## Palavras-chave

Galerias Inteligentes. Galerias de Design. Computação Gráfica. Aprendizagem. Posicionamento de Câmera. Algoritmos Genéticos. Máquinas de Suporte Vetorial (SVM).

## Abstract

Vieira, Thales; Lewiner, Thomas; Peixoto, Adelailson; Velho, Luiz.  
**Intelligent Galleries and Camera Positioning Optimization.**  
Rio de Janeiro, 2010. 78p. Tese de Doutorado — Departamento de  
Matemática, Pontifícia Universidade Católica do Rio de Janeiro.

In Visual Computing applications, several problems arise that involve a large number of parameters and have a strong subjective component. An example is the photographer problem, which tries to optimize the position, orientation and field of view to capture a tridimensional scene. Design Galleries are commonly used to support the user in such multi-parameter optimization. However, repeatedly choosing parameters for several instances of a problem turns usual Design Galleries labourious and does not use previous user interactions to help him. This work presents Intelligent Galleries, a learning approach for subjective problems such as the camera placement. The interaction of the user with a design gallery teaches a statistical learning machine. The trained machine can then imitate the user, either by classifying parameters or by automatically searching the best one. The learning process relies on a Support Vector Machines for classifying views from a collection of descriptors. We perform experiments with the automatic camera placement, which demonstrate that the proposed technique is efficient and handles scenes with occlusion and high depth complexities. This work also includes user validations of the intelligent gallery interface.

## Keywords

Intelligent Galleries. Design Galleries. Computer Graphics. Learning. Camera Positioning. Genetic Algorithms. Support Vector Machines (SVM).

# Sumário

Sumário das notações	<b>13</b>
1 Introdução	<b>15</b>
1.1 Motivação	15
1.2 Trabalhos relacionados às Galerias de Design	16
1.3 Trabalhos relacionados à otimização de posicionamento de câmera	18
2 Visão Geral das Galerias Inteligentes	<b>20</b>
2.1 Galerias de Design	20
2.2 Galerias Inteligentes	21
2.3 Aplicabilidade	23
3 Aprendizagem	<b>26</b>
3.1 Máquinas de aprendizagem	26
3.2 Usos das máquinas de aprendizagem nas Galerias Inteligentes	28
3.3 Componentes das Galerias Inteligentes	30
4 Análise	<b>33</b>
4.1 Análise heurística da influência de descritores	33
4.2 Normalização	36
4.3 Otimização da interface	37
5 Descritores para posicionamento de câmera	<b>39</b>
5.1 Extração eficiente dos descritores	40
5.2 Descritores do espaço do objeto	41
5.3 Descritores do espaço da imagem	45
6 Resultados	<b>47</b>
6.1 Capacidade de aprendizagem e reprodução de preferências	47
6.2 Eficiência da interface das Galerias Inteligentes	54
6.3 Tempos de execução	56
7 Limitações	<b>59</b>
7.1 Dimensão do espaço de parâmetros	59
7.2 Restrição a instâncias similares do problema	60
8 Conclusão	<b>61</b>
Referências Bibliográficas	<b>64</b>
A Máquinas de Suporte Vetorial	<b>69</b>
A.1 Classificador linear	69
A.2 Classificador linear com margem flexível	72
A.3 Classificador não-linear	73

A.4	Adequação às Galerias Inteligentes	74
B	Algoritmos Genéticos	<b>76</b>
B.1	Algoritmo genético usual	76
B.2	Algoritmo genético para otimização de parâmetros	77



## Lista de figuras

- 1.1 Exemplos de Galerias de Design para posicionar luz em cenas tridimensionais (esquerda); e para recolorir imagens (direita) 16
- 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  $\Omega$  que refinam as seleções do usuário. 21
- 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. 22
- 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). 24
- 2.4 Exemplo do uso das Galerias Inteligentes: a partir de uma galeria inicial (superior esquerda), o usuário seleciona bons e maus parâmetros (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). 25
- 3.1 Classificação de parâmetros: o usuário pode classificar parâmetros como bons (verdes), ruins (vermelhos) ou ignorá-los. 28
- 3.2 Galeria ordenada pela função classificadora do melhor para o pior parâmetro, começando do canto superior esquerdo no sentido anti-horário. 32
- 4.1 Exemplos de gráficos  $g_l^\Gamma$  de influência de descritores: descritor de área projetada na imagem apresentando alta influência, e descritor de médias saliências apresentando baixa influência em um banco de dados obtido através de experiências com um usuário para o problema de otimização de posicionamento de câmera. 34
- 4.2 Influência de descritores com  $\sigma = 1$  (acima) e com  $\sigma = 8$  (abaixo). A otimização do parâmetro  $\sigma$  do *kernel* Gaussiano maximiza a influência dos descritores. 35
- 4.3 Exemplo ilustrando a organização da interface das Galerias Inteligentes: no centro, em alta resolução, o visualizador permite ao observador examinar mais detalhadamente os parâmetros do mosaico. 37
- 4.4 Exemplo de galeria com excesso de parâmetros. 38
- 5.1 Visão traseira da cabeça do David: rica em alta curvatura mas pobre em saliências. 42

5.2	Descritores geométricos do espaço do objeto: curvatura gaussiana (esquerda) e saliência (direita).	43
5.3	Cena rica em oclusão: o dinossauro se esconde atrás das árvores.	44
5.4	Descritores de visibilidade de partes do espaço do objeto: <i>patches</i> (esquerda) e grupos (direita)	44
5.5	Descritores do espaço da imagem: coelho com silhueta complexa (esquerda) e vaca bem orientada em relação aos eixos da imagem (direita).	46
6.1	Posicionamento automático de câmera usando um treino voltado para as visões laterais de uma vaca (esquerda): a máquina é capaz de reproduzir com sucesso visões laterais em outros animais.	48
6.2	Análise da influência dos descritores no treino lateral da vaca (figura 6.1): a região superior direita representa os descritores mais influentes, onde destacam-se a área projetada e as altas saliências.	48
6.3	Resultado das seleções de visões de um <i>designer</i> gráfico (esquerda) e de um especialista em dinâmica de fluidos (direita).	49
6.4	Comparação entre as melhores visões escolhidas pelos profissionais em um quadro da simulação de fluidos e as visões resultantes do procedimento de seleção automática em um outro quadro da simulação: a máquina imita com sucesso as preferências do <i>designer</i> gráfico (imagens à esquerda) e do especialista em dinâmica de fluidos (imagens à direita), selecionando visões semelhantes às selecionadas por cada profissional.	50
6.5	Visões preferidas em um treino iniciado na cabeça do David (esquerda) e completada com a cabeça do Max Planck (direita).	50
6.6	Comparação de nosso método de posicionamento automática de câmera (coluna da direita) usando os dados de treino da figura 6.5 com outros trabalhos recentes: maximização das saliências (23) (superior esquerda), maximização da área projetada (34) (meio esquerda) e maximização do comprimento da silhueta (34) (inferior esquerda).	51
6.7	Análise da influência dos descritores no treino frontal duplo do David e do Max Planck (figura 6.5): seis descritores se destacam na região superior direita, com destaque para curvaturas médias e baixas, e saliências médias.	52
6.8	Treino do nó trifólio (esquerda) e posicionamento automático da câmera no nó 7.7c (direita): neste treino, foram selecionadas visões descritivas do nó, convergindo para a visão da esquerda. A melhor visão do nó 7.7c segundo a máquina reproduz a preferência do usuário por visões descritivas (direita).	53
6.9	Classificação coerente de visões do nó 7.7c (ordem crescente de qualidade de visão da esquerda para direita).	53
6.10	Treino em cena de floresta com grande potencial de oclusão, na tentativa de visualizar o dinossauro. Melhor visão do usuário (esquerda) e visão de cima da floresta, com a respectiva posição e direção da câmera (direita).	54

6.11	Resultados da seleção automática da melhor visão na cena da floresta: a máquina teve sucesso no posicionamento da câmera em outras florestas, gerando visões onde era possível visualizar o dinossauro.	55
6.12	Regressão linear do gráfico do descritor de área relativa referente ao treino do dinossauro (figura 6.10): alta correlação linear.	56
6.13	Curvas de aprendizagem de três usuários: o tempo de interação com a interface das Galerias Inteligentes diminui com o tempo, em contraste com interfaces convencionais, onde o tempo de interação permanece constante.	57
7.1	Restrição a instâncias similares: o treino do Max Planck e David não obtém resultados significativos quando aplicado ao nó (esquerda), enquanto o treino do nó também obtém um resultado sem significado no David.	60
8.1	Quadros treinados da cena de animação. As visões que focam nos grupos são selecionadas como boas.	62
8.2	Visões selecionadas automaticamente nos outros quadros chave da animação: a máquina é capaz de selecionar visões relevantes da animação.	63
A.1	Exemplo de hiperplano de margem máxima separando duas classes.	70
A.2	Formulação das Máquinas de Suporte Vetorial usando margem flexível: a condição do hiperplano ótimo é relaxada, permitindo que algumas amostras ultrapassem a margem.	73

## Lista de tabelas

- |     |  |    |
|-----|--|----|
| 5.1 | Propriedades e descritores: cada descritor está associado a uma propriedade dos vértices da cena. Estas propriedades dos vértices são quantizadas usando uma determinada quantidade de bits e codificadas em cores de 32 bits no formato RGBA. | 41 |
| 6.1 | Estabilidade da função classificadora: baixo índice de erros.  | 52 |
| 6.2 | Eficiência da interface das Galerias Inteligentes. As linhas horizontais dividem treinos de diferentes usuários.   | 58 |

## Sumário das notações

Símbolo	Significado
$\Omega$	Espaço de parâmetros
$n$	Dimensão do espaço de parâmetros $\Omega$
$p_k$	Dimensão do espaço do descritor $k$
$p$	Dimensão do espaço de descritores concatenados
$m$	Quantidade de descritores
$\Psi$	Conjunto de amostras no espaço de parâmetros
$\omega_i$	Componente do parâmetro $i$
$\Pi$	Espaço de resultados
$\pi$	Funções geradoras de resultados
$\varphi_j$	Descritor $j$
$\varphi$	Aplicação que concatena descritores
$x_i$	Vetor de descritores $\varphi(\omega_i)$
$c_i$	Classificação do parâmetro $\omega_i$
$\Gamma$	Banco de dados de treinamento
$q_v$	Ponto $v$ de um objeto
$\mathcal{L}$	Operador de Laplace-Beltrami discreto
$\kappa_H$	Operador de Curvatura Média
$\mathcal{S}$	Operador de Saliência
$\mathcal{V}_i$	Conjunto de pixels não nulos da imagem $\pi(\omega_i)$
$pix$	Pixel de uma imagem
$K$	<i>Kernel</i>
$w$	Vetor normal do hiperplano classificador das Máquinas de Suporte Vetorial
$b$	Vetor de deslocamento do hiperplano classificador das Máquinas de Suporte Vetorial
$\alpha_i$	Vetor de suporte das Máquinas de Suporte Vetorial
$\xi_i$	Variável <i>slack</i> das Máquinas de Suporte Vetorial
$\Phi$	Conjunto de descritores de parâmetros
$\chi$	Função de normalização de descritores
$x_i^l$	Coordenada $l$ do vetor de descritores $x_i = \varphi(\omega_i)$
$g_l$	Gráfico bidimensional de pontos do tipo $(x_i^l, \hat{f}(x_i))$
$\varphi^l$	Componente $l$ da aplicação $\varphi$
$\xi_i$	Variáveis <i>slack</i> das Máquinas de Suporte Vetorial
$C$	Peso associado à flexibilidade das Máquinas de Suporte Vetorial
$\phi$	Função não-linear do classificador não-linear
$\mathcal{F}$	<i>Feature space</i> das Máquinas de Suporte Vetorial
$\hat{f}$	Função linear em $\mathcal{F}$
$\nu$	Quantidade de parâmetros selecionados em uma galeria para reprodução genética
$w^p$	Posição da câmera
$w^d$	Direção de visão da câmera
$w^o$	Orientação da câmera

## Prefácio

Comecei a trabalhar na área de Computação Visual em 2005, quando ingressei no Programa de Mestrado em Matemática da Universidade Federal em Alagoas, logo após receber o grau de Bacharel em Ciência da Computação. Durante meu mestrado, fui orientado pelo Prof. Dr. Adailson Peixoto, que me direcionou para a área de Modelagem Geométrica. Nesta fase, estudei o problema de Registro de Superfícies, o qual foi tema de minha dissertação intitulada “Registro Automático de Superfícies Usando Spin-Images” (47).

Em seguida, optei por ingressar no Programa de Pós-Graduação do Departamento de Matemática da Pontifícia Universidade Católica do Rio de Janeiro para realizar meu doutoramento, sob orientação do Prof. Dr. Thomas Lewiner. Tive a oportunidade de participar de dois projetos de pesquisa para a Petrobras, com meu colega Alex Bordignon, intitulados NetGoCad e Kernelsis-br. Algumas técnicas que usamos neste trabalho foram aplicadas no projeto Kernelsis-br na área de análise das Máquinas de Suporte Vetorial.

Na área acadêmica, continuei trabalhando com Registro e Reconstrução de Superfícies, tendo publicado dois trabalhos em conferências da área, intitulados “*An iterative framework for registration with reconstruction*” (50) e “*Geometry super-resolution by example*” (48), além de um trabalho sobre união de bolas intitulado “*Scale-Space for Union of 3D Balls*” (6).

Esta tese é fruto de um trabalho realizado, desde o começo de meu doutoramento em Matemática, três anos atrás, com colaborações de meu colega Alex Bordignon, que resultou na publicação do artigo “*Learning Good Views Through Intelligent Galleries*” (49), o qual apresentamos no *Eurographics 2009* e foi publicado no *Computer Graphics Forum*.