

VISGRAF Webinar



# Computação Visual Omnidirecional

## fundamentos, desafios e aplicações

Thiago L. T. da Silveira  
tltsilveira@inf.ufrgs.br

- Prof. Dr. Thiago L. T. da Silveira

- Formação



- Atuação Profissional



- Introdução
  - O que são imagens em 360°?
  - Onde imagens em 360° se aplicam?
  - Imagens em 360° diferem das “tradicionais”?
- Modelo de Imageamento Esférico
  - Como modelar uma câmera esférica?
  - Captura de informação via imageamento esférico
- Aquisição e Representação
  - Sistemas catadióptricos, polidíoptricos e baseados em lentes *fisheye*
  - Representações equirretangular, em *cube-map* e outros

- Principais Desafios
  - O dilema do FoV: contexto *versus* distorção
  - Desafios no formato equirretangular
  - Desafios em outros formatos de representação
  - Distorções visuais e *benchmarking*
- Aplicações Seleccionadas
  - Alinhamento de imagem
  - Estimação de *layout*
  - Estimação de profundidade
- Considerações Finais

# Introdução

# O que são imagens em 360°?

- Imagens em 360° também são conhecidas como omnidirecionais, esféricas ou panorâmicas<sup>16</sup>;
- Elas armazenam a informação de toda a cena<sup>12</sup>;
- Dispositivos para captura e visualização de mídias em 360° estão se popularizando;
- Mídias em 360° aproximam o modelo ideal de imageamento (plenóptico)  $P(x, y, z, \theta, \varphi, \lambda, t)$  com três/quatro variáveis<sup>59</sup>.



# Onde imagens em 360° se aplicam?

- Panoramas são aplicáveis em virtualmente qualquer tarefa de computação visual...  
... mas são especialmente importantes em:
  - navegação de robôs<sup>16,48</sup>;
  - vigilância<sup>8,26</sup>;
  - sistemas de direção autônomos<sup>16,41</sup>;
  - experiências imersivas em AR/MR/VR<sup>2,36</sup>;
  - etc.



# Imagens em 360° diferem das “tradicionais”?

- Imagens em 360° têm um campo de visão (FoV) completo<sup>4,12</sup>;
- Elas são definidas na superfície de uma esfera<sup>3,4</sup>;
- Imagens em 360° *não* são “imagens 3D”!
- Então... como podemos processar sinais *esféricos*?



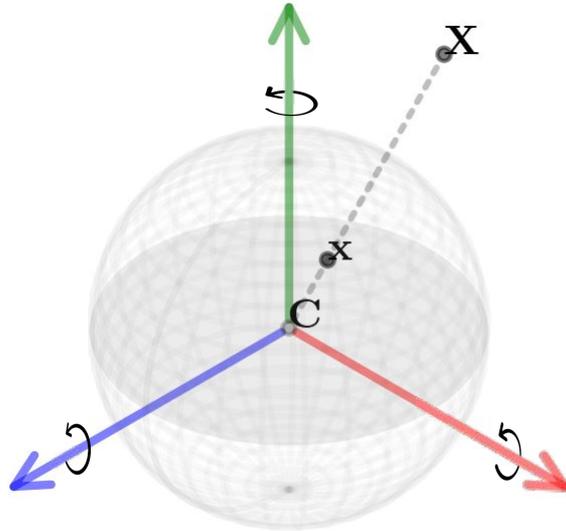
# Modelo de Imageamento Esférico

# Como modelar uma câmera esférica?

- O modelo de câmera esférica deriva das projeções centrais e esféricas<sup>11</sup>;
- A câmera é abstraída como uma esfera que “enxerga” o mundo 3D ao seu redor<sup>12,13</sup>;
- O modelo de imageamento esférico descreve uma câmera apenas com parâmetros extrínsecos (seis graus de liberdade – 6DoF)<sup>14,15</sup>;

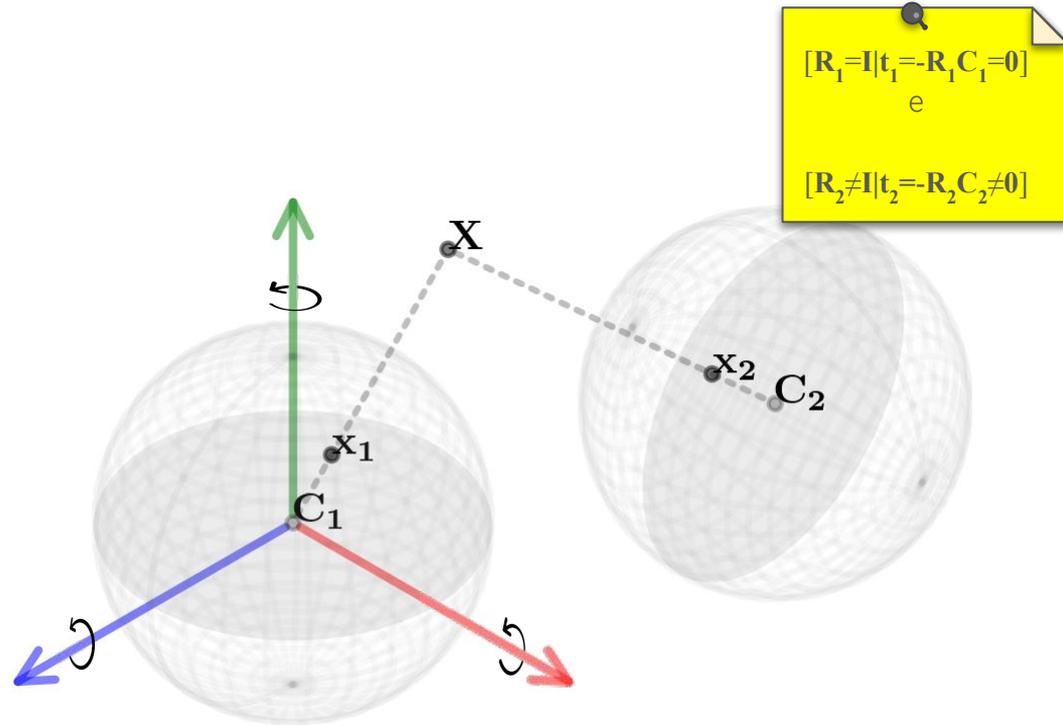
# Captura de informação via imageamento esférico

- Modelamos a câmera como uma esfera unitária;
- A câmera é centrada em  $\mathbf{C} \in \mathbb{R}^3$  e pode estar rotacionada por  $\mathbf{R} \in SO(3)$ ;
- Se  $\mathbf{C} = \mathbf{0}$  e  $\mathbf{R} = \mathbf{I}$ , então um ponto  $\mathbf{X} \in \mathbb{R}^3$  é projetado na câmera por  $\mathbf{x} = \mathbf{X}/\|\mathbf{X}\|_2$ .



# Captura de informação via imageamento esférico

- A câmera pode ter outros extrínsecos;
- $\mathbf{X}$  é projetado à  $j$ -ésima câmera ( $[\mathbf{R}_j | \mathbf{t}_j]$ ) usando  $\mathbf{x} = (\mathbf{R}_j \mathbf{X} + \mathbf{t}_j) / \|\mathbf{R}_j \mathbf{X} + \mathbf{t}_j\|_2$ , onde  $\mathbf{t}_j = -\mathbf{R}_j \mathbf{C}_j \in \mathbb{R}^3$  é chamado de vetor de translação<sup>3,4</sup>.



# Aquisição e Representação

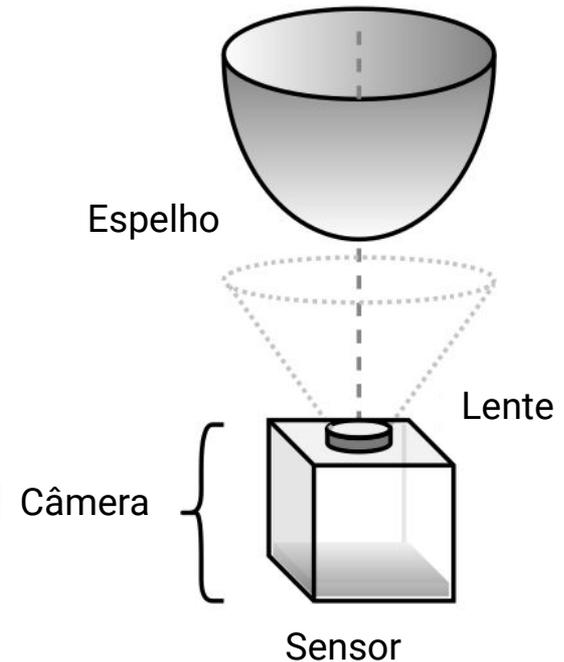
# Sistemas de captura de imagens em 360°

- Não existe um sensor que “enxerga” toda a cena<sup>17</sup>;
- Sistemas de captura existentes usam sensores planares de silício e componentes óticos<sup>16,19-21</sup>;
- A modelagem esférica é feita por *software*<sup>19</sup>;
- Há três sistemas de captura principais...



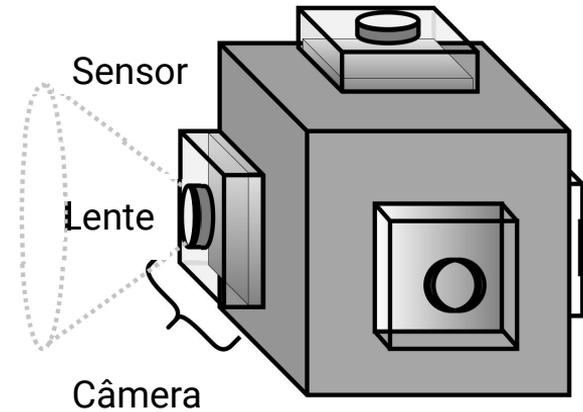
# Sistemas de captura catadióptricos

- Sistemas catadióptricos usam uma câmera regular e espelhos convexos<sup>18</sup>;
- Eles sofrem de auto-oclusão e não cobrem o FoV completo<sup>16</sup>;
- Eles são volumosos e frágeis<sup>18</sup>;
- Eles podem produzir imagens cilíndricas<sup>8</sup>.



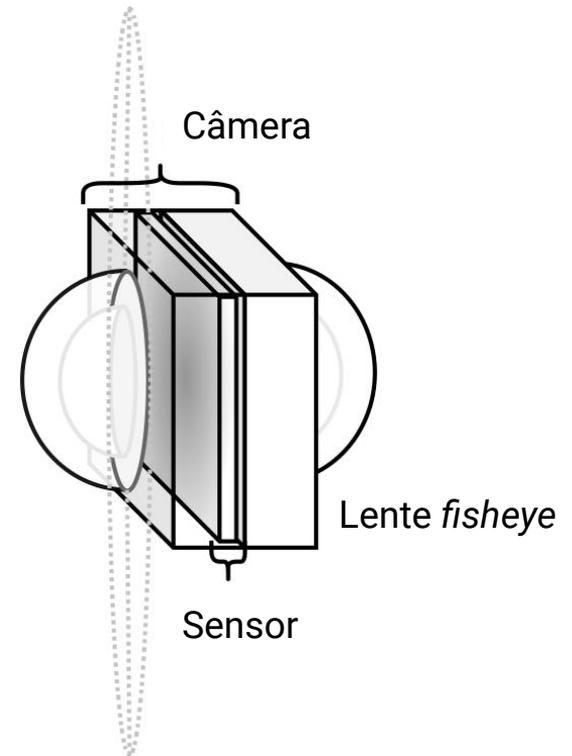
# Sistemas de captura polidióptricos

- Sistemas polidióptricos usam múltiplas câmeras em um *rig* apontando para fora<sup>16</sup>;
- Cada câmera captura uma imagem com FoV pequeno, depois combinadas via *stitching*<sup>19,20</sup>;
- Eles são volumosos e caros<sup>16</sup>;
- Eles produzem imagens de alta resolução<sup>20</sup>.



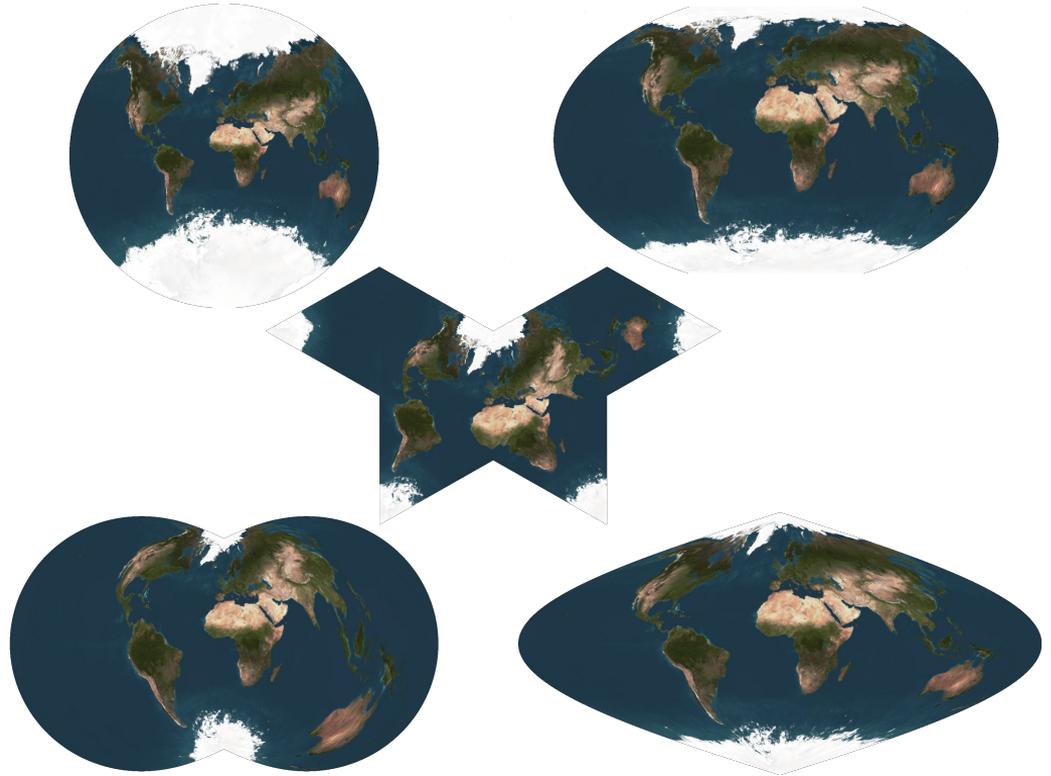
# Sistemas de captura baseados em *fisheye*

- Sistemas baseados em *fisheye* são equipados com duas lentes, uma oposta a outra, de amplo FoV<sup>21</sup>;
- Cada câmera captura uma imagem hemisférica, depois combinadas via *stitching*<sup>22</sup>;
- Eles são portáteis e baratos<sup>16,23</sup>;
- Eles produzem imagens de consumidor final<sup>23</sup>.



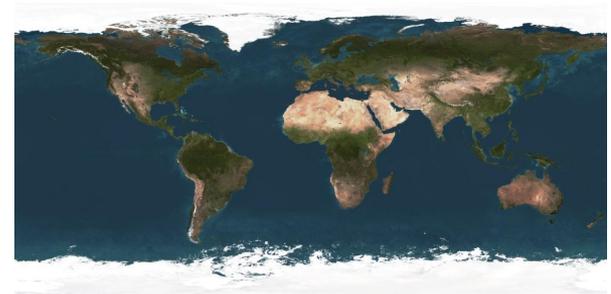
# Representações planares de imagens em 360°

- Os sinais capturados são mapeados para esfera?
- Muitos mapeamentos esfera-plano existem<sup>6</sup>;
- Cada função de mapeamento preserva propriedades específicas<sup>26</sup>;
- Todas introduzem deformações<sup>6,16</sup>.



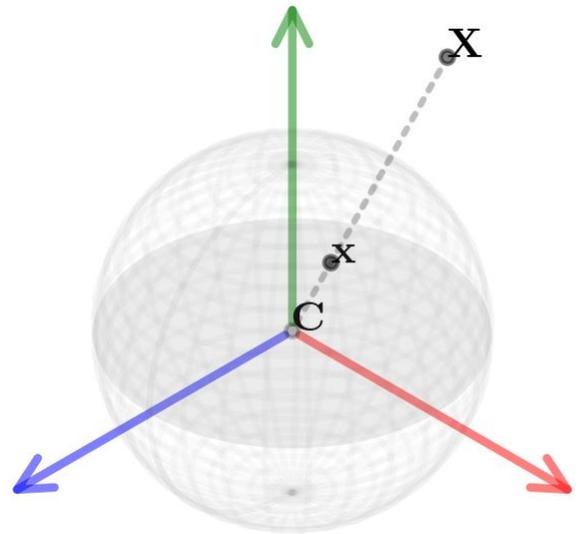
# A projeção equirretangular

- A projeção equirretangular (ERP) é o mapeamento esfera-plano padrão<sup>12</sup>;
  - Conhecida como mapeamento latitude-longitude<sup>25</sup>;
- A comunidade científica e indústria empregam o formato ERP<sup>24</sup>;
- O formato ERP é direto, mas gera uma amostragem não-uniforme no plano<sup>9,12</sup>.



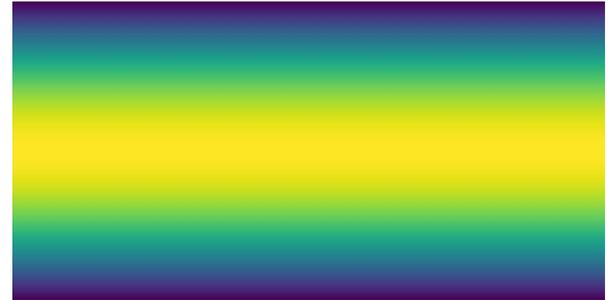
# A projeção equirretangular

- Um ponto imageado  $\mathbf{x} \in \mathcal{S}^2 \subset \mathbb{R}^3$  está na superfície da esfera <sup>12</sup>;
- Podemos reescrever  $\mathbf{x}$  usando coordenadas esféricas como  $\mathbf{x} = [r \cos\theta \sin\varphi \quad r \sin\theta \sin\varphi \quad r \cos\varphi]^\top$ , onde  $\theta \in [0, 2\pi)$  e  $\varphi \in [0, \pi)$  <sup>13</sup>;
- Bom...  $r = 1$ .



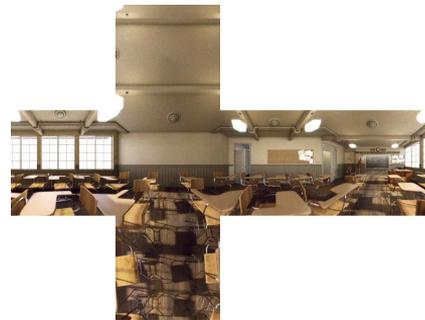
# A projeção equirretangular

- A imagem completa pode ser expressa em um mapa  $[0, 2\pi) \times [0, \pi)$ <sup>16</sup>;
- Uma imagem ERP com  $h \times w$  pixels, guarda, no pixel  $(y, x)$ , informação de<sup>13</sup>  
$$x = \lfloor L(\theta w)/(2\pi) \rfloor$$
 e  $y = \lfloor L(\phi h)/\pi \rfloor$ ;
- A amostragem horizontal em imagens ERP é proporcional a  $\sin\phi$ <sup>8,26,33</sup>.



# A projeção em *cube-map*

- A projeção em *cube-map* (CMP) é a representação multi-plano mais difundida<sup>7,27,28</sup>;
  - Conhecida como *sky-box*<sup>29</sup>;
- A CMP projeta a esfera em seis planos tangentes com FoV de  $90^\circ$ <sup>1,26</sup>;
- A CMP sofre de dependência interplano<sup>41</sup>.



# Principais Desafios

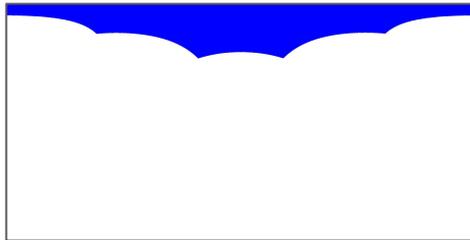
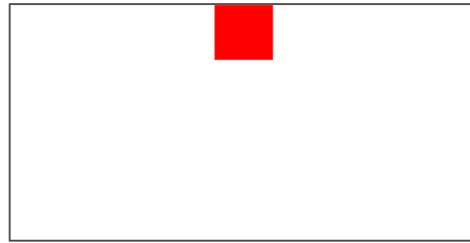
# Contexto *versus* distorção

- Quanto maior o FoV, mais informação contextual<sup>12,27</sup>; 👍
- Mas..., quanto maior o FoV, mais distorções também<sup>16,41</sup>! 🖐️
- As distorções são um problema, e se sim, como tratá-las? 🤔



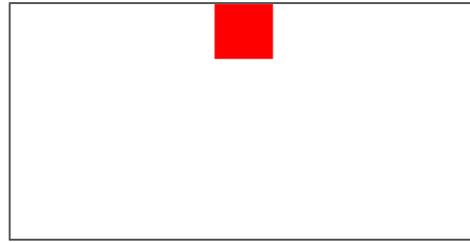
# Desafio no formato ERP

- Imagens regulares são processadas com filtros de suporte espacialmente invariante<sup>34</sup>;
- ERPs são amostradas não-uniformemente<sup>8,26</sup>;
- ERPs têm propriedade circular<sup>16,31</sup>.



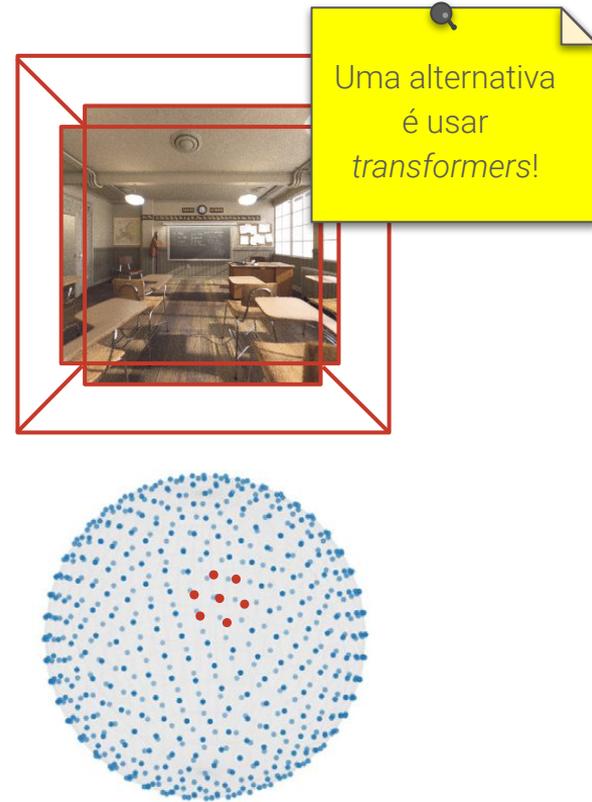
# Desafio no formato ERP

- Imagens regulares são processadas com filtros de suporte espacialmente invariante<sup>34</sup>;
- ERPs são amostradas não-uniformemente<sup>8,26</sup>;
- ERPs têm propriedade circular<sup>16,31</sup>.



# Desafio em outros formatos de representação

- Formatos multi-plano reduzem distorções mas requerem processamento interplano<sup>27,41</sup>;
- Processamento de sinais definidos em grafos não têm problemas de topologia<sup>16,44</sup>;
- No final das contas... processamento não-convencional é custoso<sup>16,40,43</sup>!



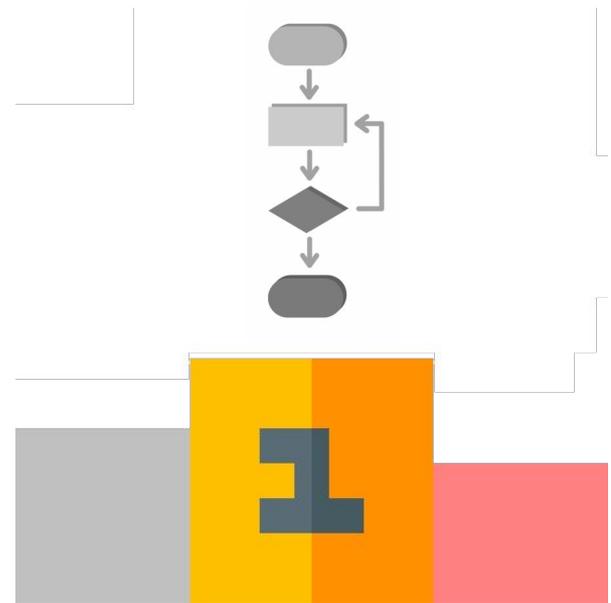
# Um desafio grande: distorções visuais

- A troca de representações leva a perda de informação<sup>7</sup>;
- Linhas de costura podem ser visíveis<sup>22</sup>;
- Objetos têm aparência variável dependendo da localização na imagem<sup>8</sup>;
- Imagens mal alinhadas podem prejudicar o processamento posterior<sup>31,40</sup>.



# Outro grande desafio: *benchmarking*

- Existem vários datasets de imagens regulares<sup>34</sup>;
- Computação visual em 360° ainda é um campo de pesquisa novo\*<sup>16</sup> ...



# Aplicações Seleccionadas

# Quais aplicações se beneficiam do FoV completo?

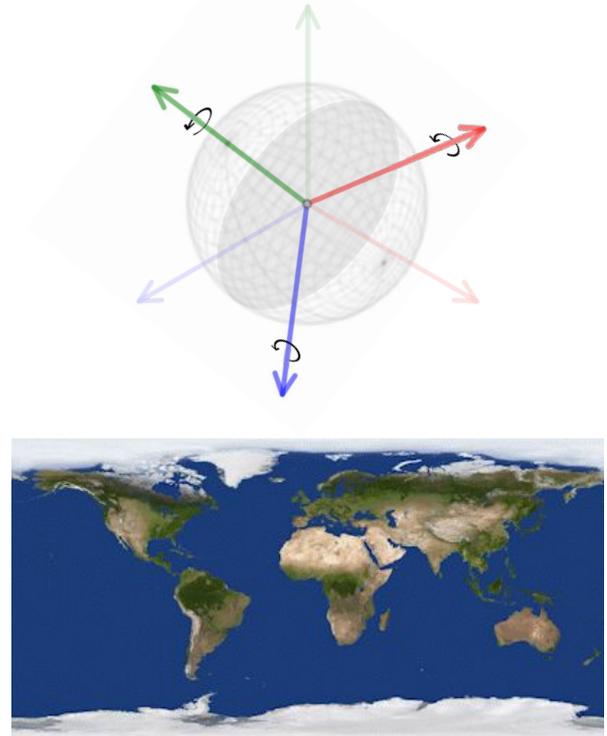
- Qualquer aplicação de computação visual pode se beneficiar de 360°;
- Discutiremos três tarefas que suportam aplicações de usuário final;
  - Alinhamento de imagem
    - Reduz mal-estar em usuários de HMD<sup>23,46</sup>;
    - Permite estabilização de vídeo<sup>14</sup>;
  - Estimação de *layout* e profundidade
    - Permitem experiências imersivas em AR/MR/VR<sup>40,58</sup>;
    - Permitem navegação de robôs/veículos em ambientes<sup>\*38,41</sup>.

Aplicações Seleccionadas

# Alinhamento de Imagens

# Definição do problema

- Imagens em 360° têm 3-DoF<sup>43</sup>;
- Elas podem ser reorientadas sem perda de informação<sup>\*45</sup>;
- Câmeras inclinadas capturam imagens “mal-orientadas”<sup>23</sup>;
- Muitas aplicações requerem panoramas alinhados<sup>53,55</sup>.



# Exemplos reais



Exemplo de [23]



Exemplo de [43]



Exemplo de [44]



Exemplo de [48]

Aplicações Seleccionadas

# Estimação de *Layout*

# Definição do problema

- Lembra? Imagens em 360° não são 3D!
- Estimar *layout* significa recuperar uma representação 3D simplificada de cenas internas<sup>16</sup>:
  - cantos<sup>35,37</sup>;
  - paredes-teto-chão<sup>55</sup>;
- Quão complexo um *layout* pode ser?



# Exemplos reais



Exemplo de [37]



Exemplo de [35]



Exemplo de [55]



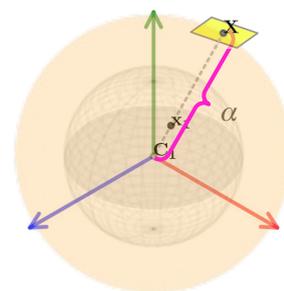
Exemplo de [54]

Aplicações Seleccionadas

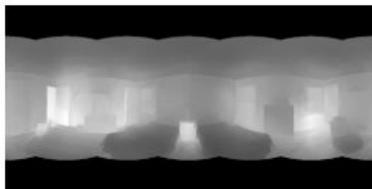
# Estimação de Profundidade

# Definição do problema

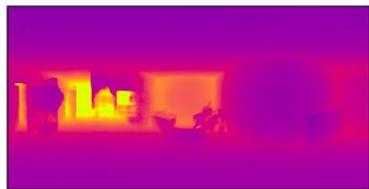
- Estimação de *layout* só se aplica em ambientes internos<sup>16</sup>;
- Estimar profundidade permite explorar aplicações completamente imersivas;
  - esparsos<sup>11,14,56</sup>/densos<sup>24,27,36,38,39,41,51,58</sup>,
  - uma<sup>27,38,39,41,51,58</sup>/duas<sup>3,11</sup>/várias<sup>15,36</sup> vistas;
- Quais condições necessárias para estimação de profundidade?



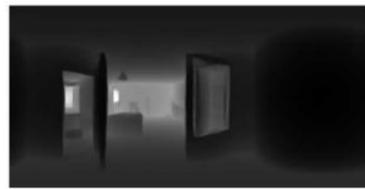
# Exemplos reais



Exemplo de [27]



Exemplo de [24]



Exemplo de [41]



Exemplo de [39]

# Considerações Finais

# O que podemos concluir?

- Imagens em 360° estão difundidas e muitas aplicações se beneficiam delas;
- Precisamos repensar vários algoritmos base para adequar à topologia de panoramas;
- A computação visual em 360° ainda sofre por falta de *benchmarking*.



Computers & Graphics 113 (2023) 89–101

Contents lists available at ScienceDirect

Computers & Graphics

Journal homepage: [www.elsevier.com/locate/cag](http://www.elsevier.com/locate/cag)

Special Issue on SIGRAPH 2022 Tutorials

**Omnidirectional visual computing: Foundations, challenges, and applications**

Thiago L.T. da Silveira<sup>a</sup>, Cláudio R. Jung

<sup>a</sup> Institute of Informatics, Federal University of Rio Grande do Sul, Porto Alegre, Brazil

ARTICLE INFO

Article history:  
Received 6 February 2023  
Received in revised form 27 April 2023  
Accepted 17 May 2023  
Available online 22 May 2023

Keywords:  
360° images  
Spherical images  
Omnidirectional images  
Panoramas

ABSTRACT

Omnidirectional images and videos are widespread due to the popularization of devices for capturing and visualization. Unlike pinhole-based imagery, omnidirectional media lie on the surface of a sphere, have a 360° × 180° field of view, and store the light intensities from an entire environment. Notably, applications involving immersive augmented, mixed, and virtual reality experiences benefit from the 360° content. Although defined on the spherical domain, omnidirectional images are often mapped to a multi-planar representation, which results in distorted images and degrades the performance of most traditional visual computing algorithms designed to work on the plane. This paper reviews the spherical camera model, the most common capturing device, and popular (multi-)planar representations of omnidirectional media. It also approaches the main challenges of omnidirectional visual computing, focusing on the deep learning paradigm, and tackles four visual computing applications that strongly explore the potential of omnidirectional imagery.

© 2023 Elsevier Ltd. All rights reserved.

**1. Introduction**

Omnidirectional images – also known as spherical, 360-degree, or panoramic images – are popular nowadays thanks to the affordability and portability of capturing devices released in the past years [1,2]. 360° images and video approximate the ideal imaging model called the plenoptic imaging model, where the whole scene visual information is captured from all possible viewpoints over time [3]. Besides classical applications, spherical media help provide immersive user experiences in novel augmented, mixed, and virtual reality (AR/MR/VR) applications when visualized in head-mounted displays (HMDs) [4]. In particular, panorama editing allows image and video manipulation and might improve users experience [5,6].

Unlike regular pinhole-based (perspective) images, which lie on the plane, omnidirectional images are defined on the surface of the unit sphere [7,8]. Spherical images have a 360° × 180° field of view (FoV) [9] that captures the light intensities from the whole scene. Fig. 1 depicts two images taken by regular and 360° cameras in a realistic 3D model<sup>1</sup>. The two cameras are placed at the same location and orientation. Remarkably, the topology of pinhole-based and 360° images differ, so traditional visual



computing algorithms designed to work on the planar domain are not directly applicable to omnidirectional media.

Although panoramic images are indeed defined on the spherical domain, they are commonly represented in a (multi-)planar form [9–11]. Many sphere-to-plane mapping functions can be used to generate the planar representation, but they all introduce distortions [12,13]. Although a panorama can be represented in the plane (like in “world map” format), the visual computing algorithm that takes it as input still needs to consider the introduced distortions to be accurate in its task [4,13].

Compared to the conventional visual computing field, omnidirectional visual computing is still embryonic, and only a few

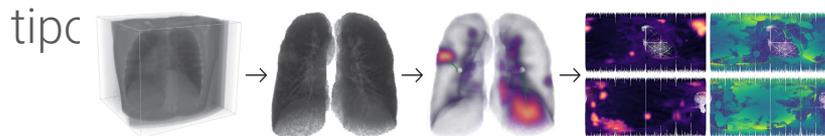
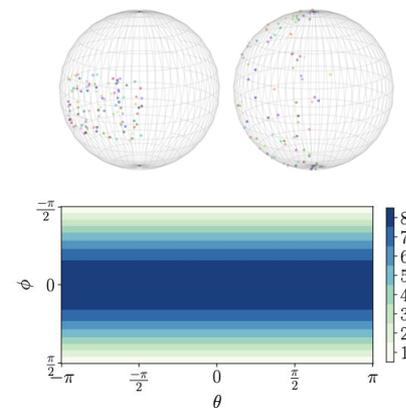
<sup>a</sup> Corresponding author.  
E-mail addresses: [thiagoltd@inf.ufrgs.br](mailto:thiagoltd@inf.ufrgs.br) (T.L.T. da Silveira), [crjung@inf.ufrgs.br](mailto:crjung@inf.ufrgs.br) (C.R. Jung).

<sup>1</sup> The Chamoon 3D computer-generated model is available under CC0 license at <https://www.chamoon.org>.

<https://doi.org/10.1016/j.cag.2023.05.011>  
0097-4497/2023 Elsevier Ltd. All rights reserved.

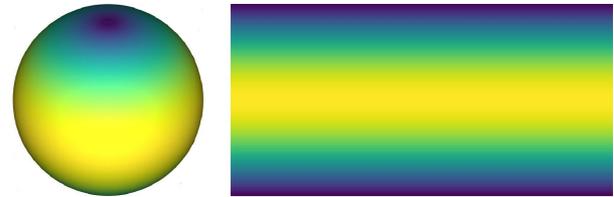
# No que tenho trabalhado?

- Alinhamento de imagens e estimação de pose de câmera com 5-DoF/6-DoF;
- Estimação de *layout* e profundidade;
- Classificação e detecção de objetos;
- Tarefas de *baixo nível*: casamento de pontos-chave, supersegmentação, compressão, etc.;
- Representações holísticas de outros de dados (e.g., CTs de pulmão)...



# Colaborações com o VISGRAF?

- Compressão de imagens?
- Síntese de vistas com suporte a 3DoF+/6-DoF?
  - Com ou sem NeRFs;
- Aplicações com representações neurais implícitas de cenas?
- O que mais?



# Referências

# Referências

- [1] J. Huang et al., “6-DoF VR videos with a single 360-camera,” in IEEE VR, 2017, pp. 37–44.
- [2] A. Serrano et al., “Motion parallax for 360° RGBD video,” IEEE TVCG, vol. 25, no. 5, pp. 1817–1827, 2019.
- [3] S. Li, “Binocular spherical stereo,” IEEE TITS, vol. 9, no. 4, pp. 589–600, 2008.
- [4] J. Fujiki et al., “Epipolar Geometry Via Rectification of Spherical Images,” in Computer Vision/Computer Graphics Collaboration Techniques. Springer Berlin Heidelberg, 2007, vol. 4418, pp. 461–471.
- [5] W. Yang et al., “Object Detection in Equirectangular Panorama,” ICPR, vol. 2018-August, pp. 2190–2195, 2018.
- [6] Y.-C. Su and K. Grauman, “Learning Spherical Convolution for Fast Features from 360° Imagery,” in NIPS, 2017, pp. 529–539.
- [7] R. G. d. A. Azevedo et al., “Visual Distortions in 360-degree Videos,” IEEE TCSVT, vol. 30, no. 8, pp. 2524–2537, 2020.
- [8] J. Cruz-Mota et al., “Scale invariant feature transform on the sphere: Theory and applications,” IJCV, vol. 98, no. 2, pp. 217–241, 2012.
- [9] T. L. T. da Silveira et al., “Fast and accurate superpixel algorithms for 360° images,” SIGPRO, vol. 189, p. 108277, 2021.
- [10] R. Hartley and A. Zisserman, Multiple View Geometry in Computer Vision. Cambridge, 2003.
- [11] S. Li and K. Fukumori, “Spherical stereo for the construction of immersive vr environment,” in IEEE VR, 2005, pp. 217–222.
- [12] T. L. T. da Silveira and C. R. Jung, “Perturbation Analysis of the 8-Point Algorithm: A Case Study for Wide FoV Cameras,” in IEEE CVPR, 2019, pp. 11 757–11 766.
- [13] T. Akihiko , “Two-and three-view geometry for spherical cameras,” Workshop on Omnidirectional Vision, Camera Networks and Non-classical Cameras, vol. 105, pp. 29–34, 2005.
- [14] H. Guan and W. A. P. Smith, “Structure-From-Motion in Spherical Video Using the von Mises-Fisher Distribution,” IEEE TIP, vol. 26, no. 2, pp. 711–723, 2017.
- [15] B. Krolla, M. Diebold, B. Goldlucke, and D. Stricker, “Spherical light fields,” BMVC, no. 67.1-67.12, 2014.
- [16] T. L. T. da Silveira, P. G. L. Pinto, J. Murrugarra-Llerena, and C. R. Jung, “3d scene geometry estimation from 360° imagery: A survey,” ACM CSUR, 2022, just Accepted.
- [17] J. D. Adarve and R. Mahony, “Spherepix: A data structure for spherical image processing,” IEEE RA-L, vol. 2, no. 2, pp. 483–490, 2017.
- [18] S. K. Nayar, “Catadioptric Omnidirectional Camera\*,” in CVPR, 1997, pp. 482–488.
- [19] S. Im et al., “Allaround depth from small motion with a spherical panoramic camera,” in ECCV, 2016, pp. 156–172.

# Referências

- [20] G. Fangi et al., “Improving spherical photogrammetry using 360° OMNI-Cameras: Use cases and new applications,” *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 42, no. 2, pp. 331–337, 2018.
- [21] Y. Shan and S. Li, “Descriptor Matching for a Discrete Spherical Image With a Convolutional Neural Network,” *IEEE Access*, vol. 6, pp. 20 748–755, 2018.
- [22] I. Lo et al., “Image stitching for dual fisheye cameras,” in *IEEE ICIP*, 2018, pp. 3164–3168.
- [23] R. Jung et al., “Deep360Up: A Deep Learning-Based Approach for Automatic VR Image Upright Adjustment,” in *IEEE VR*, 2019, pp. 1–8.
- [24] M. Eder et al., “Pano Popups: Indoor 3D Reconstruction with a Plane-Aware Network,” in *IEEE 3DV*, 2019, pp. 76–84.
- [25] C. C. Gava et al., “Dense Scene Reconstruction from Spherical Light Fields,” in *IEEE ICIP*, 2018, pp. 4178–4182.
- [26] L. S. Ferreira et al., “Local Moebius transformations applied to omnidirectional images,” *C&G*, vol. 68, pp. 77–83, 2017.
- [27] T. L. T. da Silveira et al., “Indoor Depth Estimation from Single Spherical Images,” in *IEEE ICIP*, 2018, pp. 2935–2939.
- [28] F. Dai et al., “Freely Explore the Scene with 360° Field of View,” in *IEEE VR*, 2019, pp. 888–889.
- [29] S. Song et al., “Im2Pano3D: Extrapolating 360° Structure and Semantics Beyond the Field of View,” in *IEEE/CVF CVPR*, vol. 1, 2018, pp. 3847–3856.
- [30] M. Eder et al., “Tangent images for mitigating spherical distortion,” in *IEEE CVPR*, 2020.
- [31] Y. Lee, et al., “Spherephd: Applying cnns on 360° images with non-euclidean spherical polyhedron representation,” *IEEE TPAMI*, pp. 1–1, 2020.
- [32] B. Coors et al., “SphereNet: Learning spherical representations for detection and classification in omnidirectional images,” *ECCV*, pp. 525–541, 2018.
- [33] F. De Simone et al., “Geometry-driven quantization for omnidirectional image coding,” in *PCS*, 2017.
- [34] I. Goodfellow et al., *Deep Learning*. The MIT Press, 2016.
- [35] C. Sun et al., “HorizonNet: Learning Room Layout with 1D Representation and Pano Stretch Data Augmentation,” pp. 1047–1056, 2019.
- [36] T. L. T. da Silveira and C. R. Jung, “Dense 3D Scene Reconstruction from Multiple Spherical Images for 3-DoF+ VR Applications,” in *IEEE VR 2019*, pp. 9–18.
- [37] N. Zioulis et al., “Single-shot cuboids: Geodesics-based end-to-end manhattan aligned layout estimation from spherical panoramas,” p. 104160, 2021.
- [38] N. Zioulis et al., “OmniDepth: Dense Depth Estimation for Indoors Spherical Panoramas,” in *ECCV*, 2018, pp. 453–471.
- [39] K. Tateno et al., “Distortion-Aware Convolutional Filters for Dense Prediction in Panoramic Images,” *ECCV*, pp. 732–750, 2018.

# Referências

- [40] C. Fernandez-Labradoret et al., "Corners for layout: End-to-end layout recovery from 360 images," IEEE RA-L, pp. 1–1, 2020.
- [41] F.-E. Wang et al., "Bifuse: Monocular 360 depth estimation via bi-projection fusion," in Proceedings of the IEEE/CVF CVPR, 2020.
- [42] J. Xiao et al., "Recognizing scene viewpoint using panoramic place representation," in IEEE CVPR, 2012, pp. 2695–2702.
- [43] M. A. Bergmann et al., "Gravity alignment for single panorama depth inference," in SIBGRAPI, 2021, pp. 1–8.
- [44] R. Jung et al., "Upright adjustment with graph convolutional networks," in IEEE ICIP, 2020, pp. 1058–1062.
- [45] J. Murrugarra-Llerena et al., "Pose estimation for two-view panoramas based on keypoint matching: A comparative study and critical analysis," in IEEE/CVF CVPR, June 2022, pp. 5202–5211.
- [46] J. Jung et al., "Robust upright adjustment of 360 spherical panoramas," The Visual Computer, vol. 33, no. 6, pp. 737–747, 2017.
- [47] L. Zhang et al., "Vanishing point estimation and line classification in a manhattan world with a unifying camera model," IJCV, vol. 117, no. 2, pp. 111–130, 2016.
- [48] Y. Shan and S. Li, "Discrete spherical image representation for cnnbased inclination estimation," IEEE Access, vol. 8, pp. 2008–2022, 2019.
- [49] H. Yang and H. Zhang, "Modeling room structure from indoor panorama," in ACM SIGGRAPH, 2014, pp. 47–55.
- [50] H. Jia and S. Li, "Estimating structure of indoor scene from a single full-view image," in IEEE ICRA, 2015, pp. 4851–4858.
- [51] C. Sun et al., "Hohonet: 360 indoor holistic understanding with latent horizontal features," pp. 2573–2582, 2021.
- [52] G. Pintore et al., "State-of-the-art in automatic 3d reconstruction of structured indoor environments," Computer Graphics Forum, vol. 39, no. 2, 2020.
- [53] Y. Zhang et al., "PanoContext: A whole-room 3D context model for panoramic scene understanding," in ECCV, 2014.
- [54] F.-E. Wang et al., "LED2-Net: Monocular 360° layout estimation via differentiable depth rendering," pp. 12 956–12 965, 2021.
- [55] G. Pintore et al., "AtlantaNet: Inferring the 3D indoor layout from a single 360 image beyond the Manhattan world assumption," in ECCV, 2020.
- [56] A. Pagani and D. Stricker, "Structure from Motion using full spherical panoramic cameras," in IEEE ICCV, 2011, pp. 375–382.
- [57] G. Albanis et al., "Pano3d: A holistic benchmark and a solid baseline for 360° depth estimation," in IEEE/CVF CVPR 2021, pp. 3722–3732.
- [58] H. Jiang et al., "Unifuse: Unidirectional fusion for 360° panorama depth estimation," IEEE RA-L, vol. 6, no. 2, pp. 1519–1526, 2021.

Perguntas?



CAPES FAPERGS

VISGRAF Webinar



# Computação Visual Omnidirecional

## fundamentos, desafios e aplicações

Thiago L. T. da Silveira  
tltsilveira@inf.ufrgs.br