

---

# Image Stitching

Marcos Oliveira

## O que é

- Processo de combinar várias imagens fotográficas com campos de visão sobrepostos para produzir um panorama segmentado ou uma imagem de alta resolução.
- A maioria das abordagens requerem sobreposições quase exatas entre as imagens e exposições idênticas para produzir resultados perfeitos. Também é conhecido como mosaicos.

## Etapas

- Aquisição de imagem: Requer a seleção da posição e aquisição de imagens para calibragem.
- Remapeamento de imagem (Image Warping) - É o processo de alterar a geometria da imagem deformando-a para ajustá-la às imagens adjacentes que devem contribuir para um panorama.



## Etapas

- Mistura de imagem (Blending images): Uma vez que os pixels de origem foram remapeados na superfície final do composto, ainda devemos decidir como combiná-los para criar o panorama. Esta etapa é usada para remover imperfeições visíveis como por exemplo às diferenças de exposição.



## Etapas

- Corte (Cropping): Refere-se à remoção de áreas indesejadas de uma imagem fotográfica ou ilustrada.

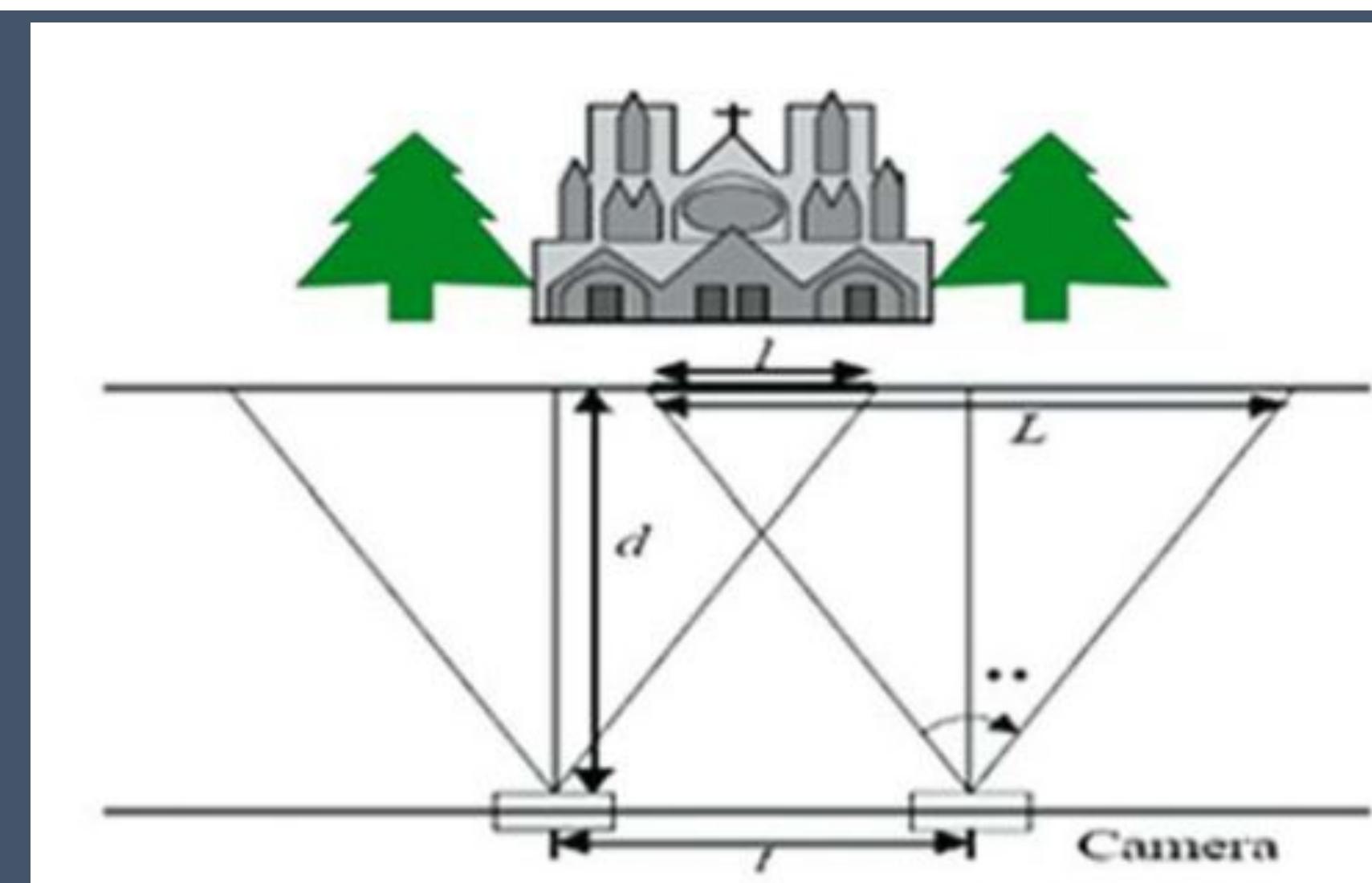
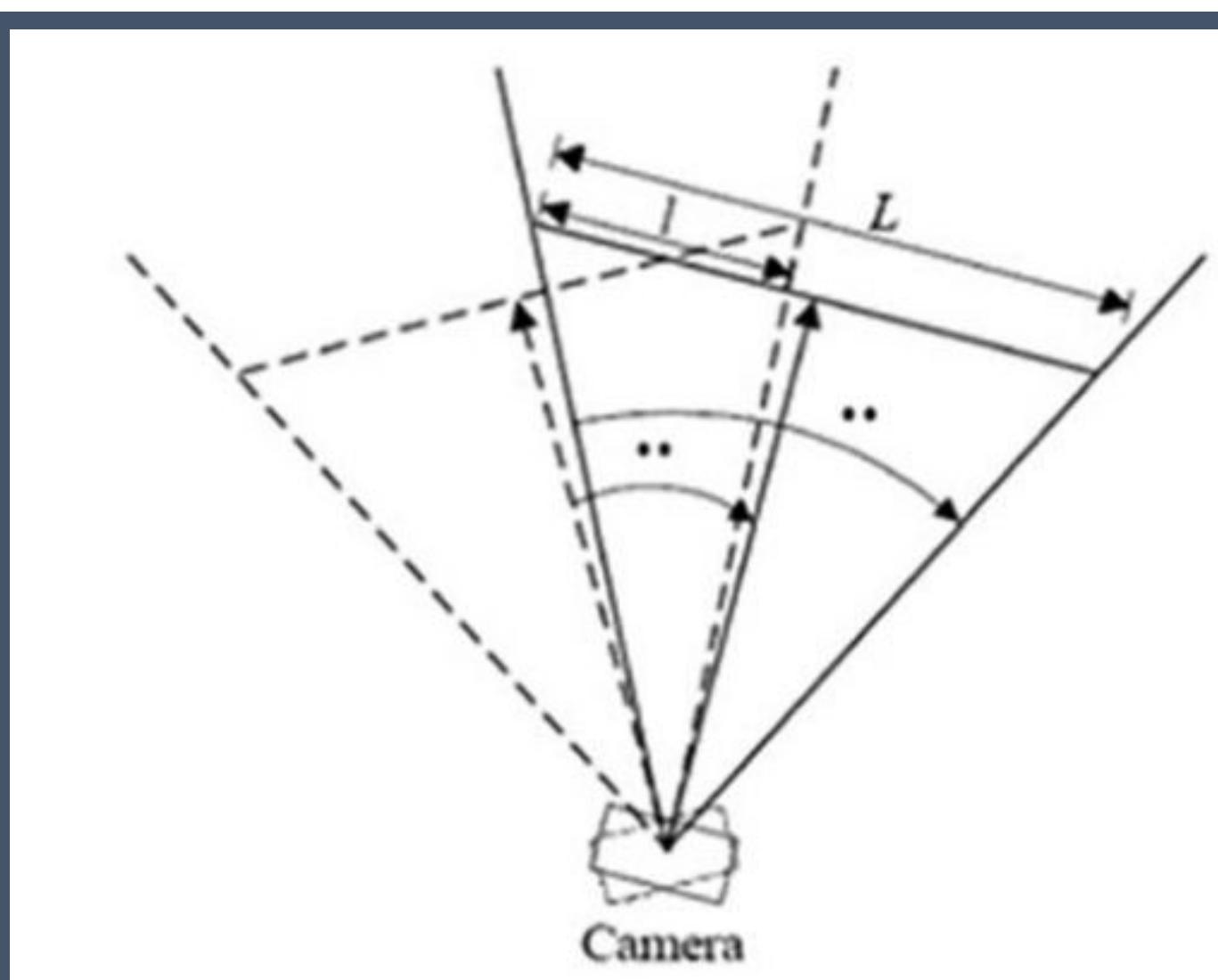
## Funcionamento Básico

- OpenCV
  - 1. Capture uma sequência de imagens
  - 2. Calcule a transformação entre a segunda imagem e a primeira imagem passado por parâmetro.
  - 3. Mude a segunda imagem para se sobrepor à primeira
  - 4. Misture os dois para criar um panorama
  - 5. Se houver mais imagens, repita.



# Aquisição

- Aquisição de imagens por rotação de camera:
  - Uma imagem é tirada a cada rotação da câmera até que a faixa desejada seja coberta.
  - Desejável que 50% da imagem se sobreponha à imagem anterior e os outros 50% da imagem se sobreponha à imagem seguinte.
  - No entanto, o tamanho real da região de sobreposição pode diferir do calculado se a câmera for inclinada, ou seja, houver rotação da câmera diferente da rotação necessária em torno do eixo Y.
- Aquisição de imagens por translação:
  - Neste método de aquisição, a câmera é deslocada em uma direção paralela ao plano de imagem por meio de uma placa deslizante.
  - É importante garantir que os planos da imagem sejam paralelos à direção das translações da câmera. Caso contrário, o tamanho dos objetos nas imagens varia conforme a câmera é deslocada, causando problemas na montagem da imagem.



# Remapeamento de imagem (Image Warping)

- Detecção de pontos-chave (keypoints)

Uma característica (ponto-chave) é uma informação significativa extraída de uma imagem que fornece uma compreensão mais detalhada da imagem. Em outras palavras, características são objetos, que devem ser distintos, salientes e de preferência espalhados pela imagem e fáceis de detectar.

Os objetos podem ser, por exemplo, pontos de referência ou áreas segmentadas e até mesmo a intensidade pode ser vista como um recurso. O número de objetos detectados comuns em ambas as imagens deve ser suficientemente alto para ser capaz de realizar uma boa estimativa do modelo de transformação.

Sift



Brisk



# Remapeamento de imagem (Image Warping)

- Detecção de pontos-chave (keypoints)
  - SIFT:
    - Scale Invariant Feature Transformation
    - SIFT tenta extrair características invariantes de escala usando uma abordagem de filtragem em estágios.
- Localidade: os recursos são locais, portanto robustos à oclusão e desordem (sem segmentação anterior)
- Distinção: características individuais podem ser combinadas a um grande banco de dados de objetos
- Quantidade: muitos recursos podem ser gerados até mesmo para objetos pequenos
- Eficiência: desempenho próximo do tempo real
- Extensibilidade: pode ser facilmente estendida a uma ampla gama de diferentes tipos de recursos, com cada um adicionando robustez.

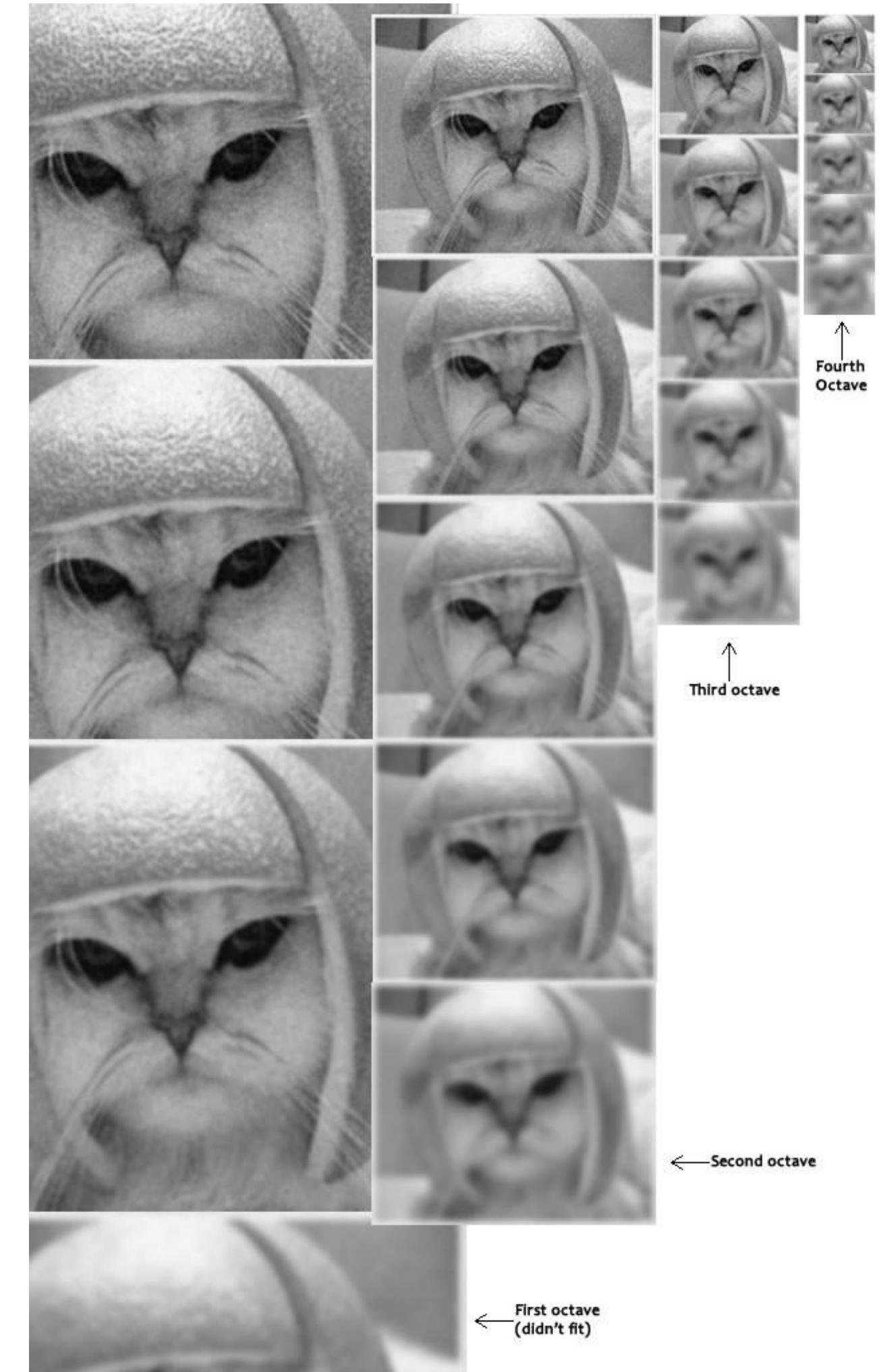
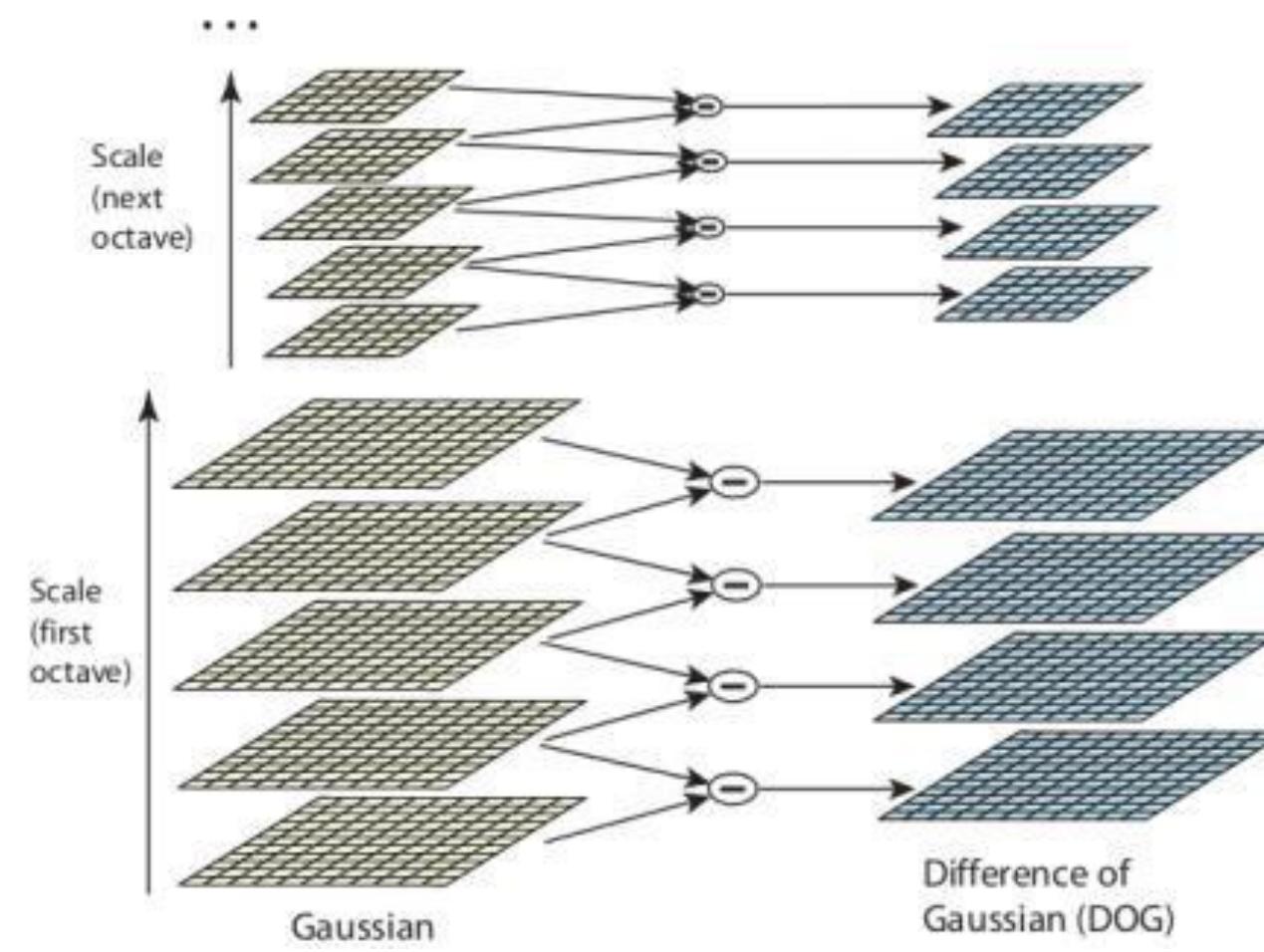
SIFT



# Remapeamento de imagem (Image Warping)

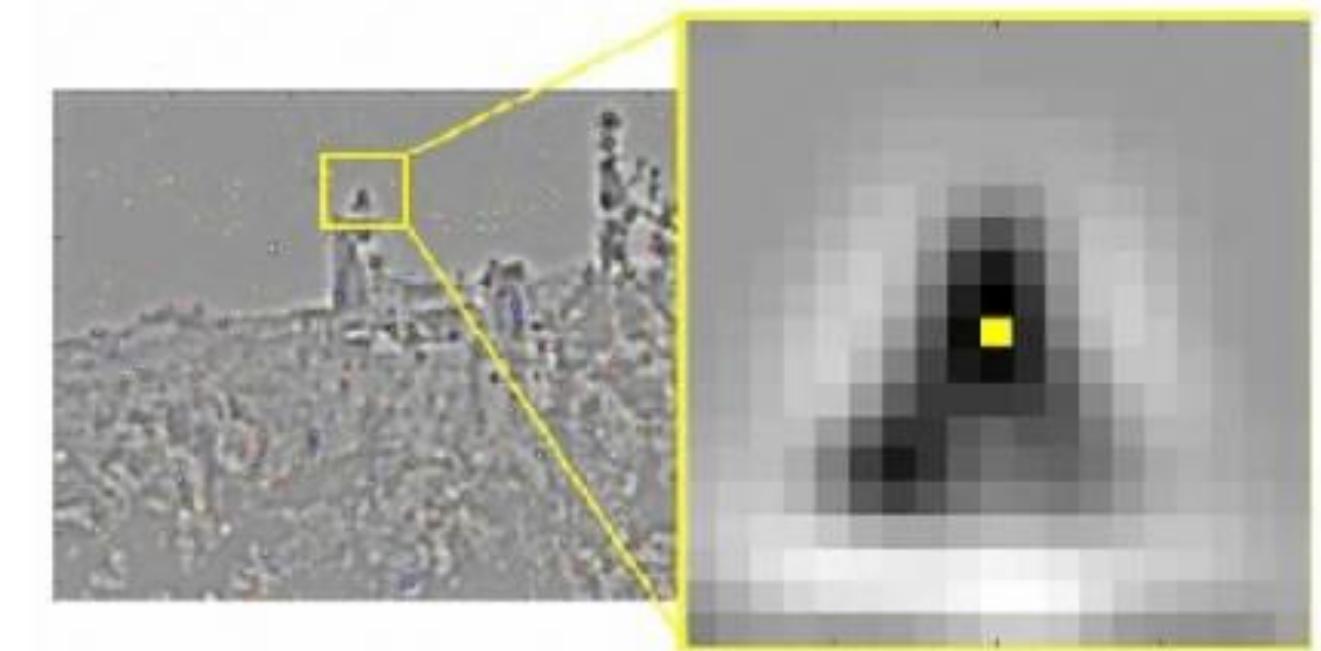
- SIFT- Funcionamento Básico:

- Seleção de pico em espaço de escala: localização potencial para encontrar recursos.
  - O espaço de escala de uma imagem é uma função  $L(x, y, \sigma)$  que é produzida a partir da convolução de um kernel gaussiano (Desfoque) em diferentes escalas com a imagem de entrada. O espaço de escala é separado em oitavas e o número de oitavas e a escala dependem do tamanho da imagem original. Portanto, geramos várias oitavas da imagem original. O tamanho da imagem de cada oitava é a metade do anterior.
  - As imagens borradadas são então utilizadas para gerar outro conjunto de imagens, a Diferença de Gaussianas (DoG), obtida como a diferença de desfoque gaussiano de uma imagem com dois  $\sigma$  diferentes, sejam  $\sigma$  e  $k\sigma$ . Este processo é feito para diferentes oitavas da imagem na Pirâmide Gaussiana.



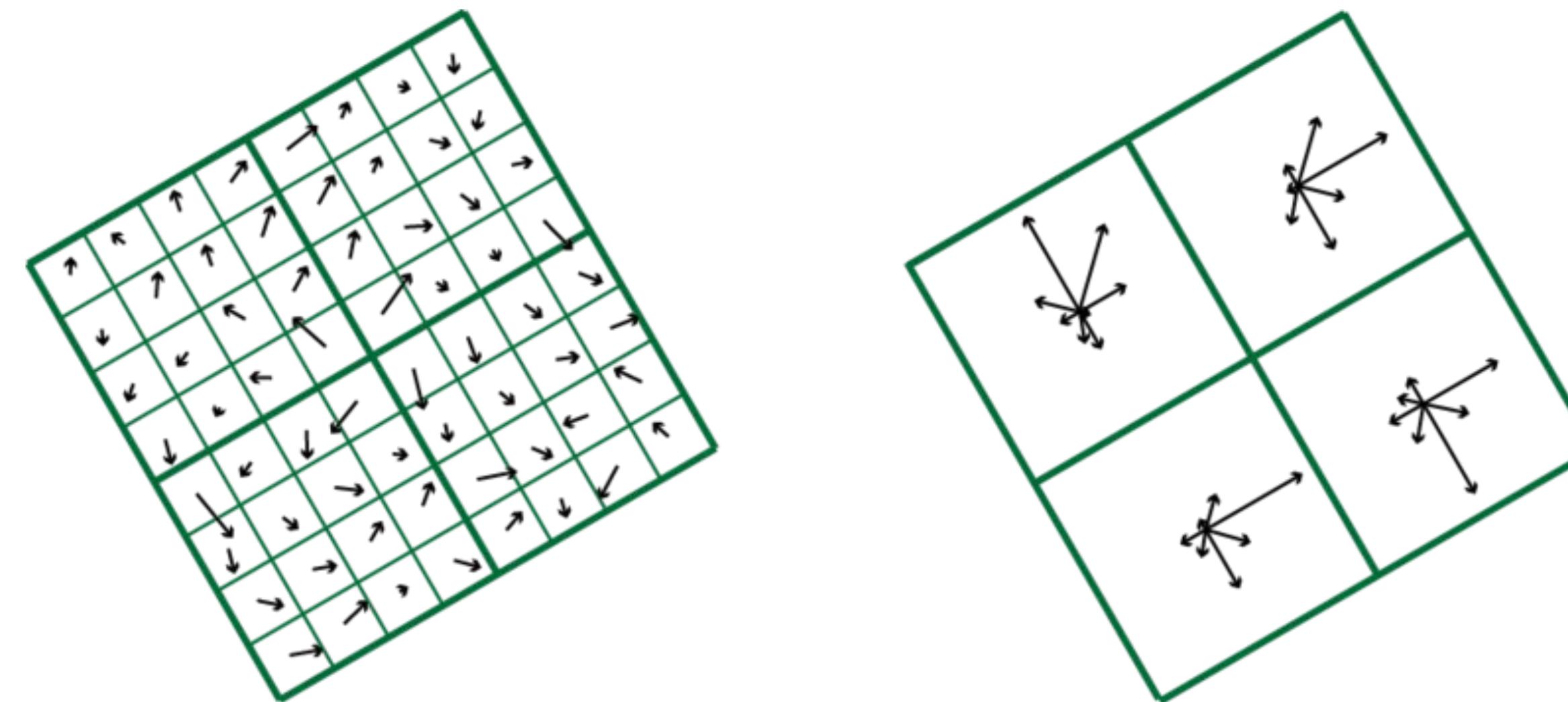
# Remapeamento de imagem (Image Warping)

- SIFT- Funcionamento Básico:
  - Localização do ponto-chave: localizar com precisão os pontos-chave do recurso.
    - Um pixel de uma imagem é comparado a seus 8 vizinhos, bem como a 9 pixels na escala seguinte e 9 pixels nas escalas anteriores somando 26 verificações
    - Se o pixel for um extremo local, é um ponto-chave potencial. Entretanto alguns deles ficam ao longo de uma borda ou não têm contraste suficiente.
    - Para solução é aplicada a expansão da série de Taylor do espaço de escala para obter uma localização mais precisa dos extremos e, se a intensidade nesses extremos for menor que um valor limite (0,03), ela será rejeitada.
  - Atribuição de orientação: Atribuição de orientação aos pontos-chave.
    - Uma vizinhança é feita em torno da localização do ponto-chave, dependendo da escala, e a magnitude e a direção do gradiente são calculadas nessa região.
    - É criado um histograma de orientação com 36 caixas cobrindo 360 graus. E a quantidade que é adicionada à caixa é proporcional à magnitude do gradiente naquele ponto.
    - O pico mais alto no histograma é obtido e qualquer pico acima de 80% dele também é considerado para calcular a orientação, criando pontos-chave com a mesma localização e escala, mas em direções diferentes. Isso contribui para a estabilidade da combinação.



## Remapeamento de imagem (Image Warping)

- SIFT- Funcionamento Básico:
  - Descritor de ponto-chave: Descrever os pontos-chave como um vetor de alta dimensão.
    - Dadas as estimativas de escala e orientação para um ponto de interesse, uma grade retangular é disposta no domínio da imagem, centrada no ponto de interesse, com sua orientação determinada pelos picos principais do histograma e com espaçamento proporcional à escala de detecção do ponto de interesse.
  - Para dar pesos mais fortes às orientações de gradiente perto do ponto de interesse, as entradas no histograma também são ponderadas por uma função de janela gaussiana centrada no ponto de interesse e com seu tamanho proporcional à escala de detecção do ponto de interesse. Tomados em conjunto, os histogramas locais calculados em todos os pontos da grade 4x4 e com 8 direções quantizadas levam a um descritor de imagem com dimensões 4x4x8 para cada ponto de interesse.



# Remapeamento de imagem (Image Warping)

Correspondência de pontos-chave -->(Raw matches (Brute force): 510)



- Pontos-chave entre duas imagens são combinados identificando seus vizinhos mais próximos geralmente.

(Na imagem exemplificado com distância euclidiana)

## Extensões

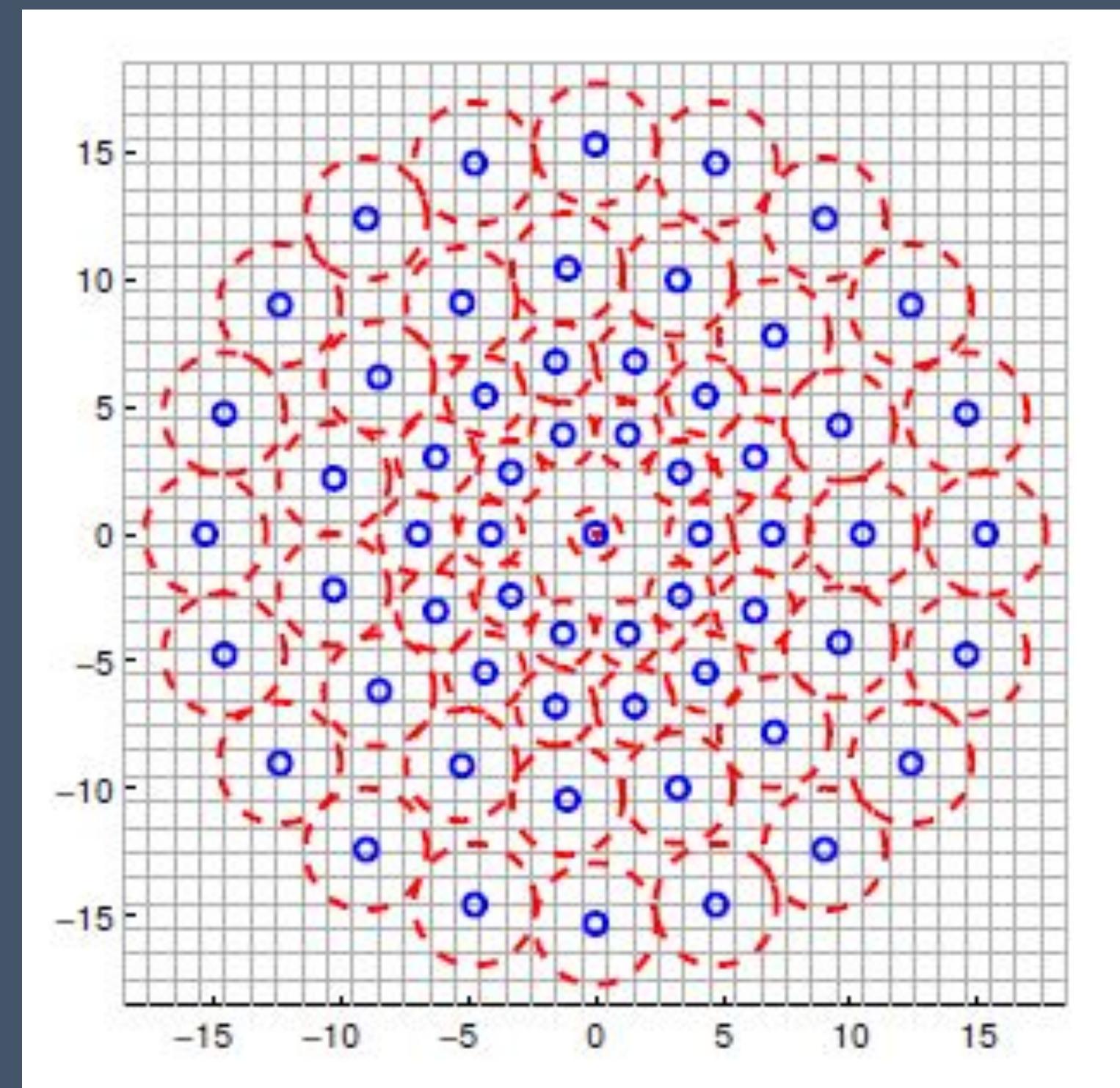
- PCA SIFT
- Colour SIFT
- SIFT-like image descriptors for spatio-temporal recognition
- Dense SIFT

# Remapeamento de imagem (Image Warping)

- Detecção de pontos-chave (keypoints)
  - BRISK:
    - Se comporta diferente do método SIFT
    - Os pixels são amostrados em anéis concêntricos.



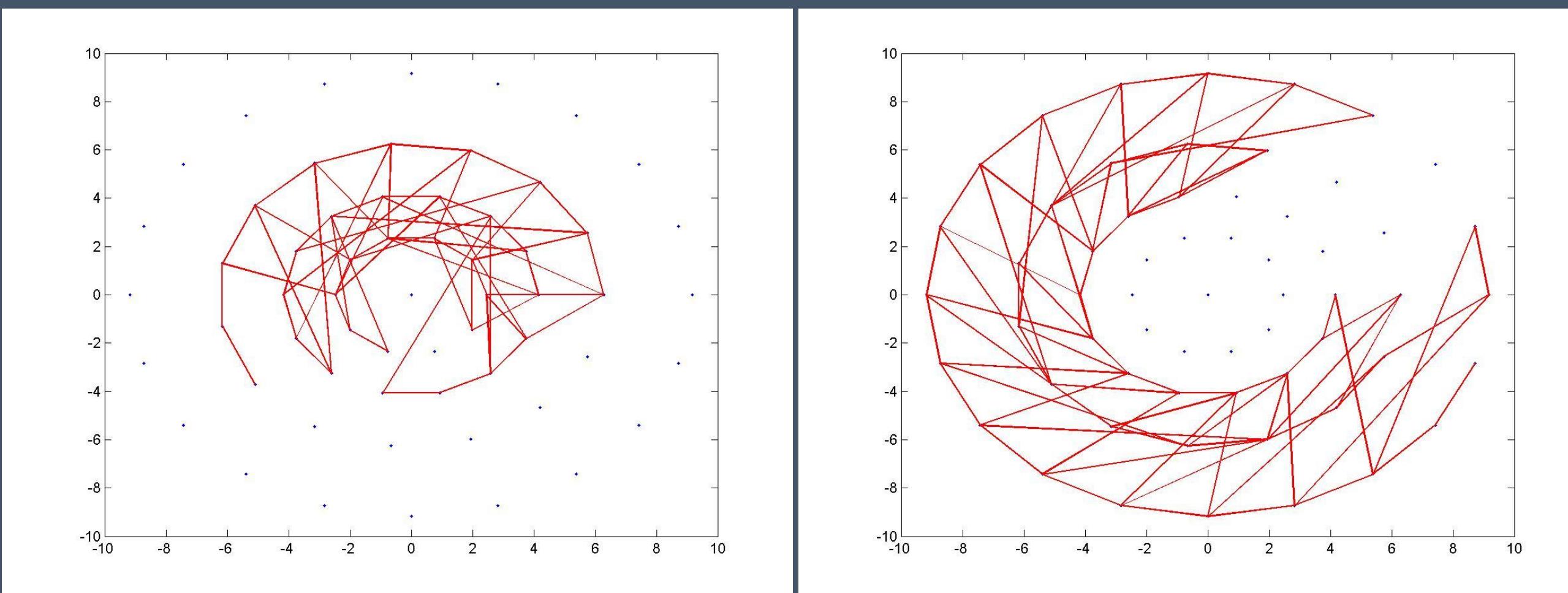
- Ao considerar cada ponto de amostragem, é capturado um pequeno patch em torno dele e aplicado a suavização gaussiana. O círculo vermelho na figura ilustra o tamanho do desvio padrão do filtro gaussiano aplicado a cada ponto de amostragem.



# Remapeamento de imagem (Image Warping)

- Detecção de pontos-chave (keypoints)
  - BRISK:
    - Ele produz dois pares. Um par de longa distância e outro significa par de curta distância. Os pares de longa distância são aqueles pares cuja distância deve ser considerada acima do valor limite (threshold) e usado para determinar a orientação e os pares curtos são usados para as comparações de intensidade que constroem o descriptor.

Correspondência de pontos-chave -->(Raw matches (Brute force): 570)



## Remapeamento de imagem (Image Warping)

---

- Combinação de pontos-chave (matching keypoints)
  - **Brute-Force Matcher**
    - O BF Matcher combina o descritor de características de uma imagem com todos os outros descritores de outra imagem e retorna a correspondência com base na distância, como distancia euclidiana (L2 Norm) utilizado pelo SIFT e a Hamming distance (L2 Hamming) utilizada pelo BRISK.
  - Pode ser utilizado apenas como ‘força bruta’
  - Pode ser utilizado com k-vizinhos próximos, o qual encontrará combinações ótimas porém não necessariamente ser as melhor

# Remapeamento de imagem (Image Warping)

- Combinação de pontos-chave (matching keypoints)
  - Brute-Force Matcher

SIFT



BF Matcher -->(Raw matches (Brute force): 510)



BF Matcher -->(Raw matches (KNN): 687)

# Remapeamento de imagem (Image Warping)

- Combinação de pontos-chave (matching keypoints)
  - Brute-Force Matcher

BRISK



BF Matcher -->(Raw matches (Brute force): 570)



BF Matcher -->(Raw matches (KNN): 874)

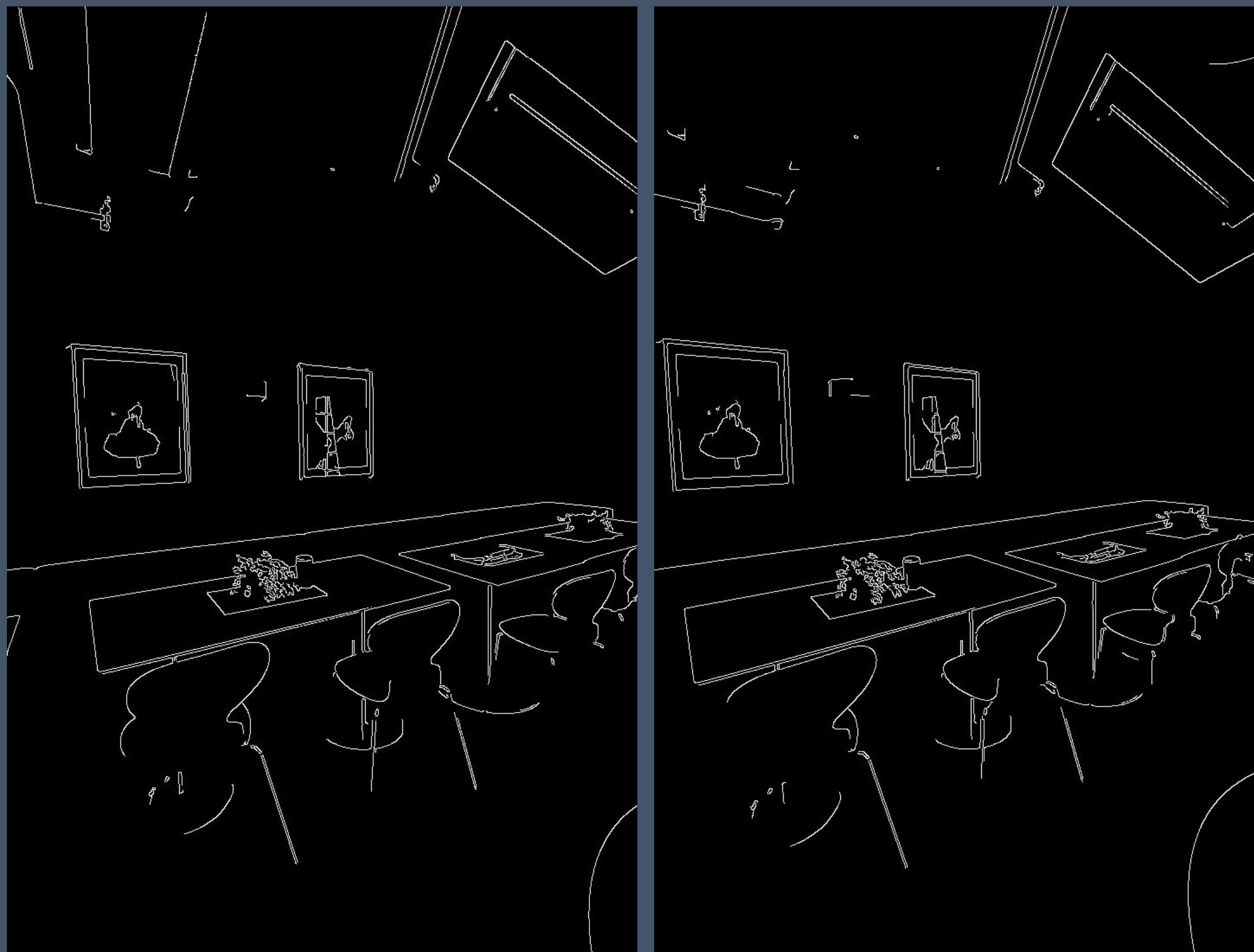
# Remapeamento de imagem (Image Warping)

- Combinação de pontos-chave (matching keypoints)
  - FLANN (Fast Approximate Nearest Neighbor Search Library)
    - Contém uma coleção de algoritmos otimizados para busca rápida do vizinho mais próximo em grandes conjuntos de dados e para recursos dimensionais elevados.
- Floresta k-d randomizada
- Árvore k-means de busca prioritária



# Remapeamento de imagem (Image Warping)

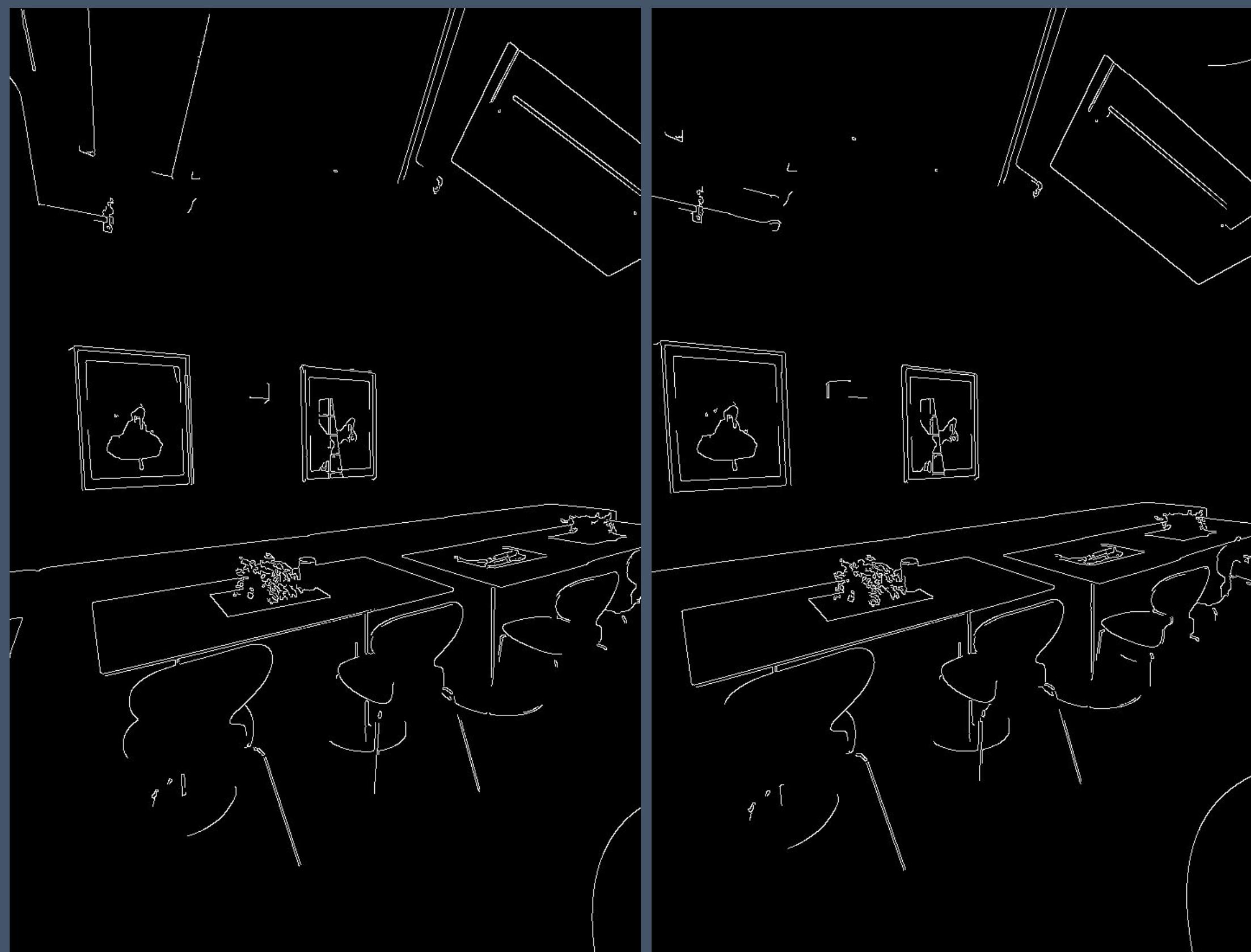
- Combinação de pontos-chave (matching keypoints)
  - FLANN (Fast Approximate Nearest Neighbor Search Library)



FLANN Matcher -->(Raw matches (KNN): 1051)

# Remapeamento de imagem (Image Warping)

- Combinação de pontos-chave (matching keypoints)
  - FLANN (Fast Approximate Nearest Neighbor Search Library)



FLANN Matcher -->(Raw matches (KNN): 3895)

# Remapeamento de imagem (Image Warping)

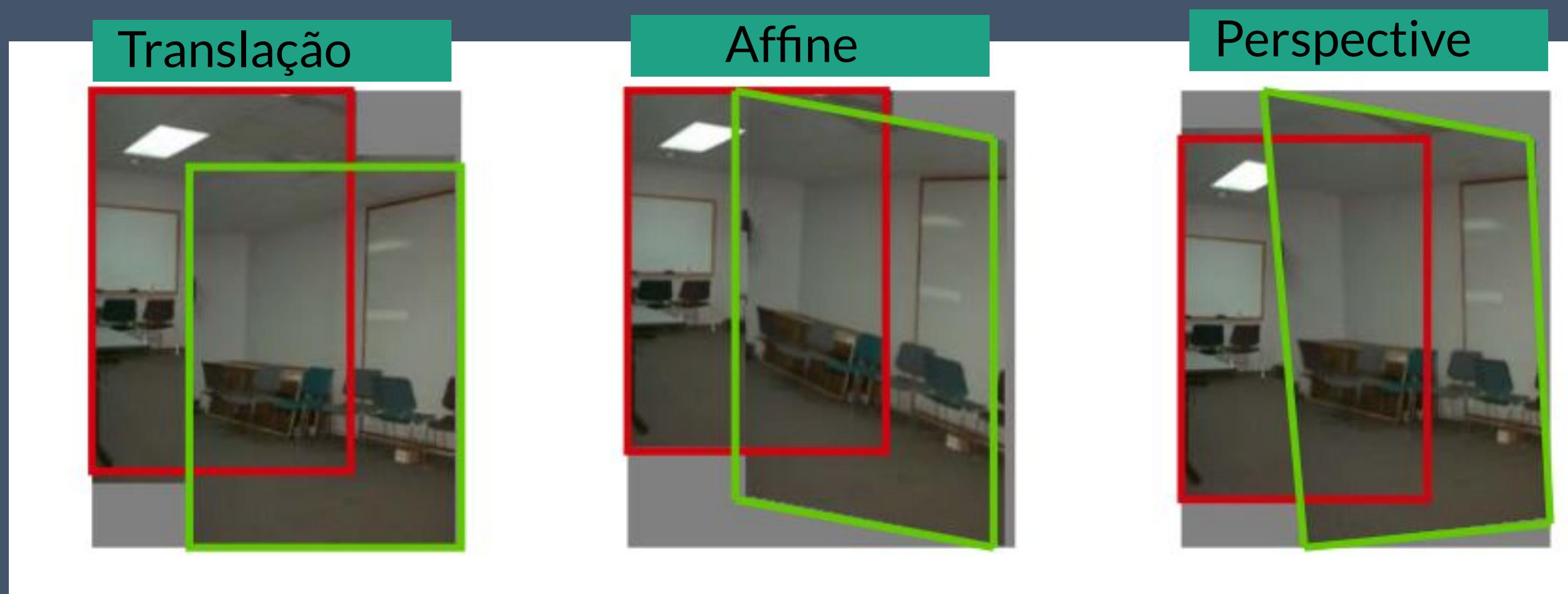
- **Image Warping**

- Translação - 2 graus de liberdade
- Affine - 6 graus de liberdade
- Perspective - 8 graus de liberdade

- Homografia (Perspective)

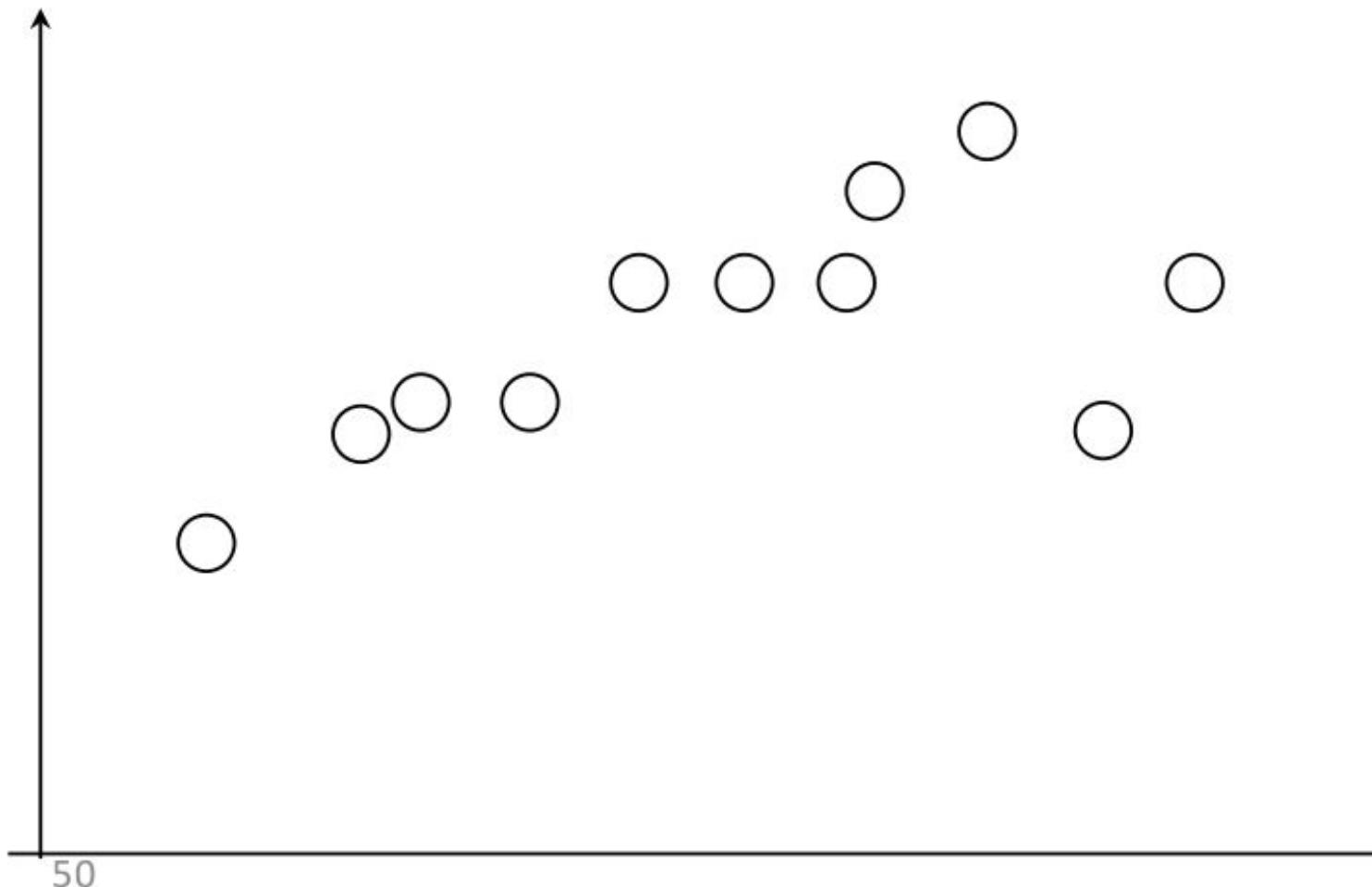
$$\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = H \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} h_{00} & h_{01} & h_{02} \\ h_{10} & h_{11} & h_{12} \\ h_{20} & h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix}$$

- Homografia (Perspective)
- RANSAC - RANdom SAmple Consensus
  - Algoritmo iterativo para ajustar modelos lineares e projetado para ser robusto para outliers.
  - Modelos como regressão linear usam estimativa de mínimos quadrados para ajustar o melhor modelo aos dados. No entanto, os mínimos quadrados comuns são muito sensíveis a outliers. Como resultado, ele pode falhar se o número de outliers for significativo. Ao contrário, o RANSAC resolve esse problema ajustando o modelo no subconjunto de pontos identificados como inliers.

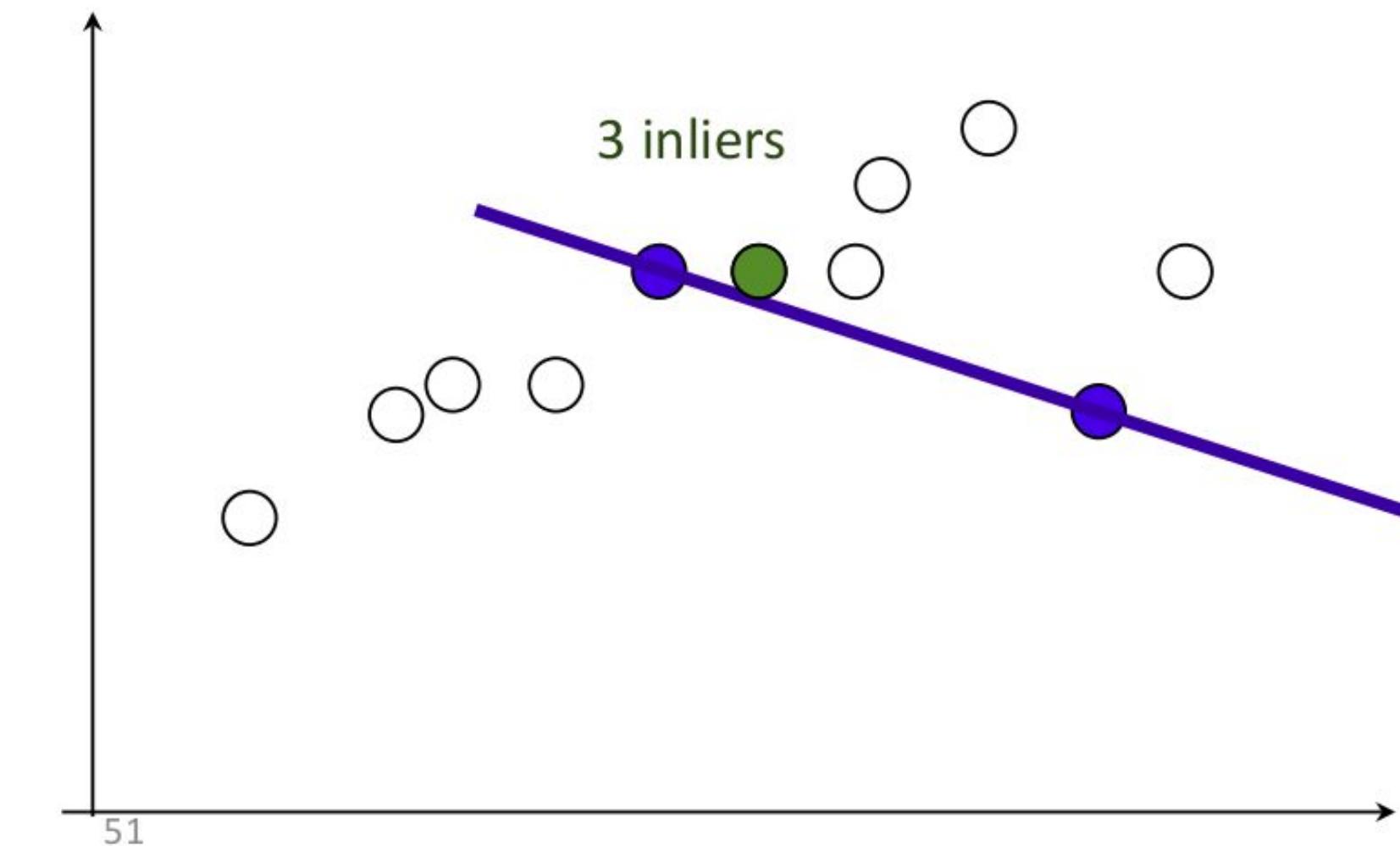


## Remapeamento de imagem (Image Warping)

- RANSAC: Exemplo básico.
  - Em vez de homografia  $H$  (8 números) ajuste  $y = ax + b$  (2 números  $a, b$ ) para pares 2D.

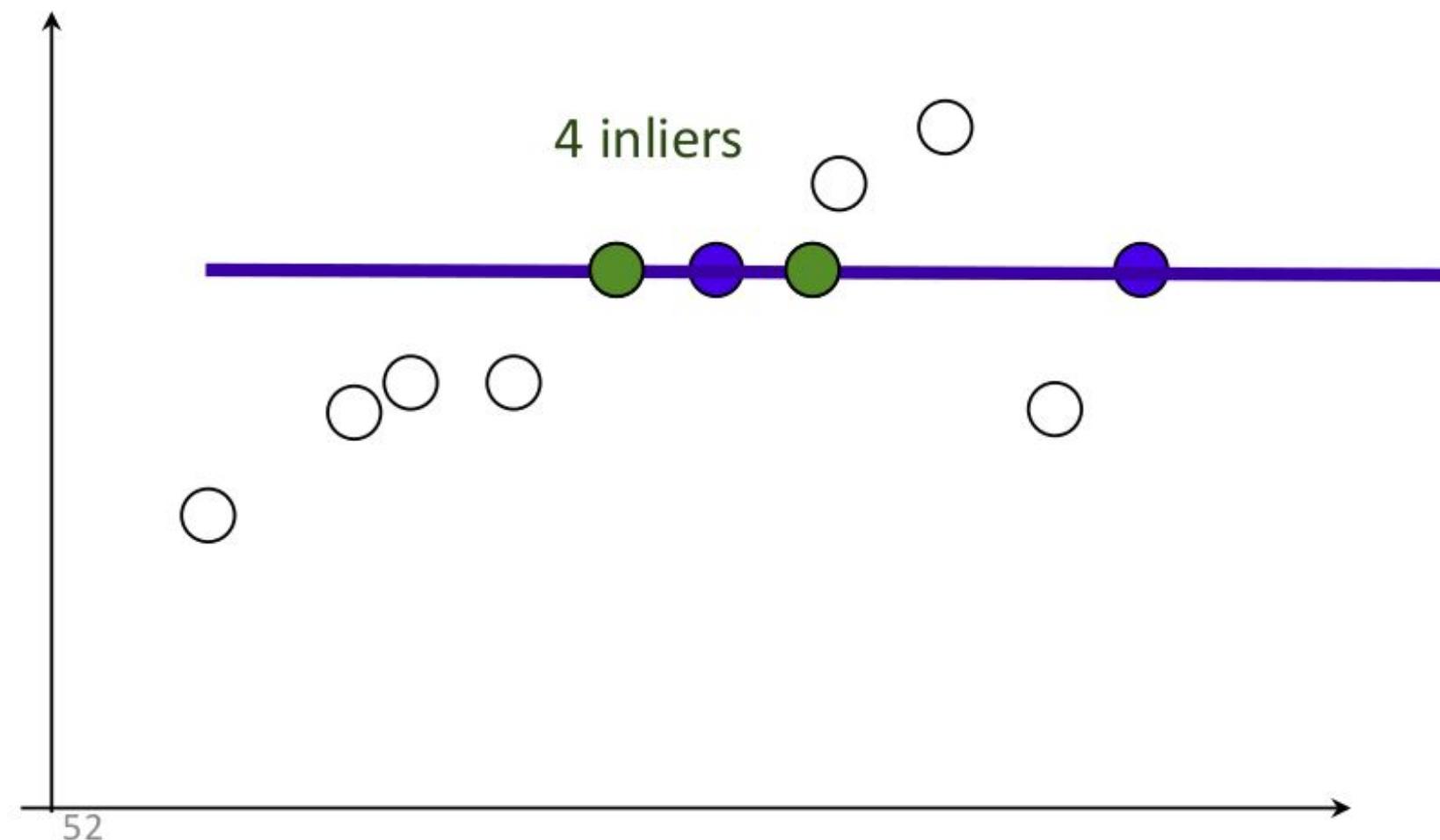


- Escolha 2 pontos
- Ajuste a linha
- Conte inliers

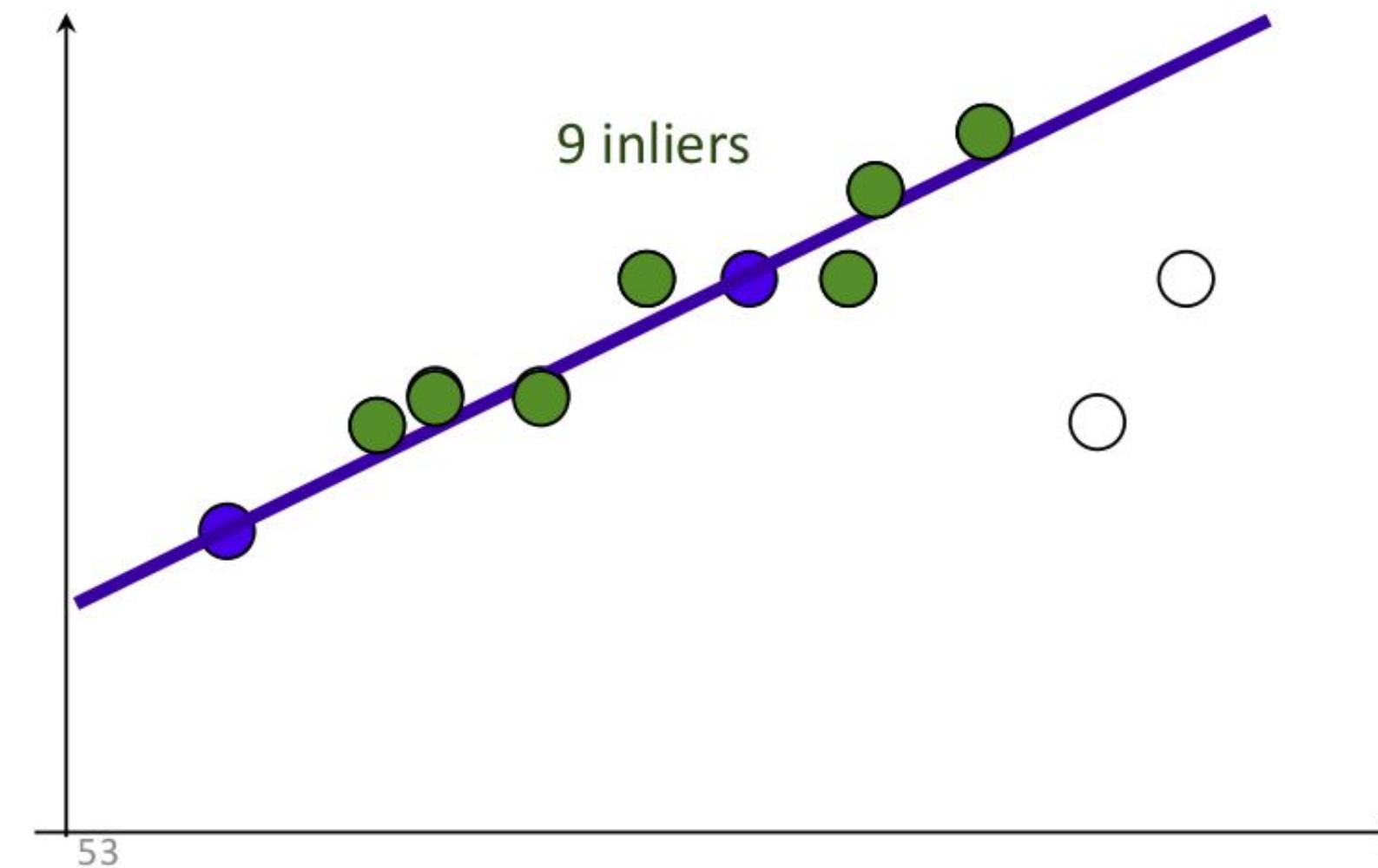


## Remapeamento de imagem (Image Warping)

- Escolha 2 pontos
- Ajuste a linha
- Conte inliers

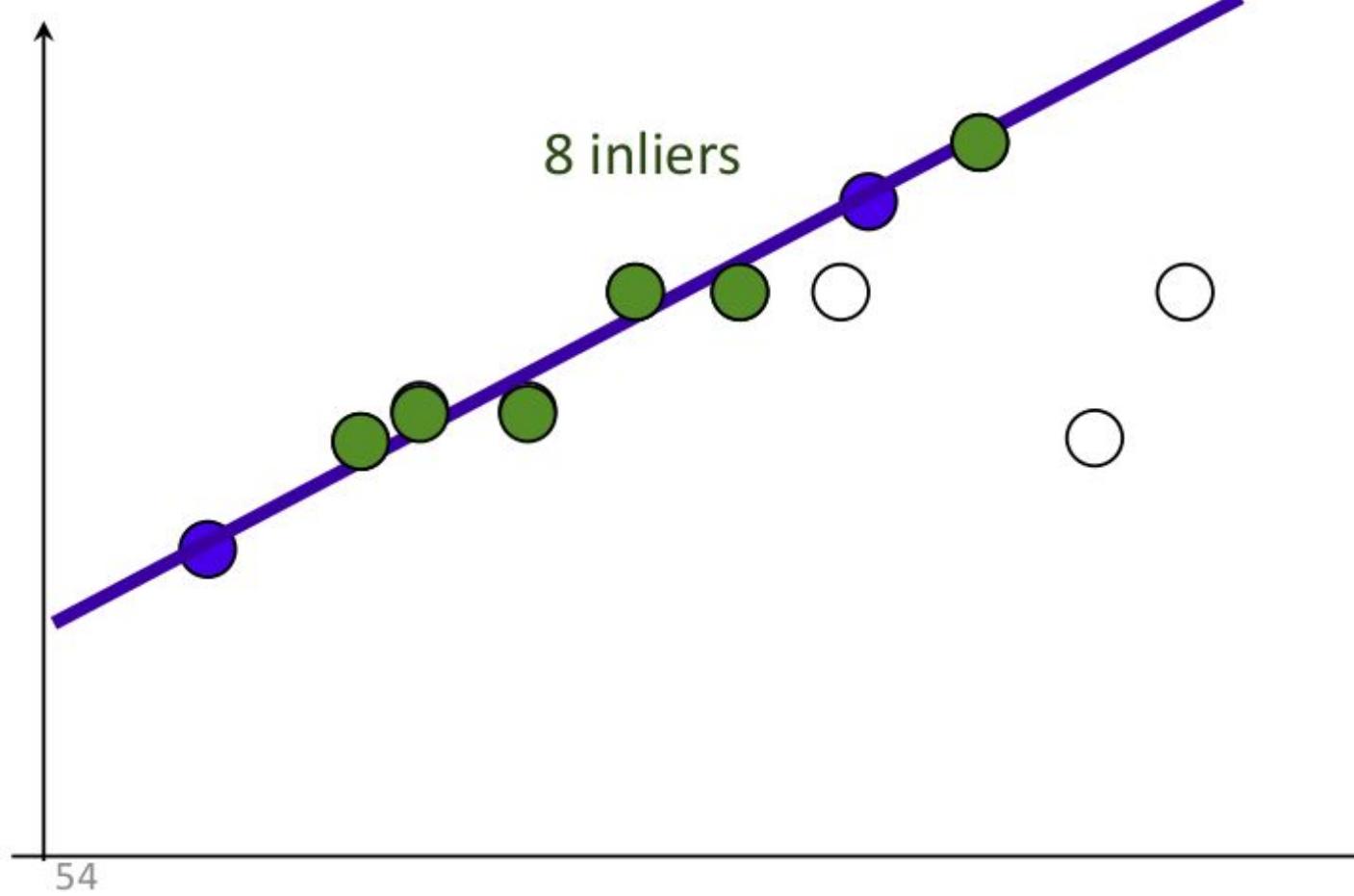


- Escolha 2 pontos
- Ajuste a linha
- Conte inliers

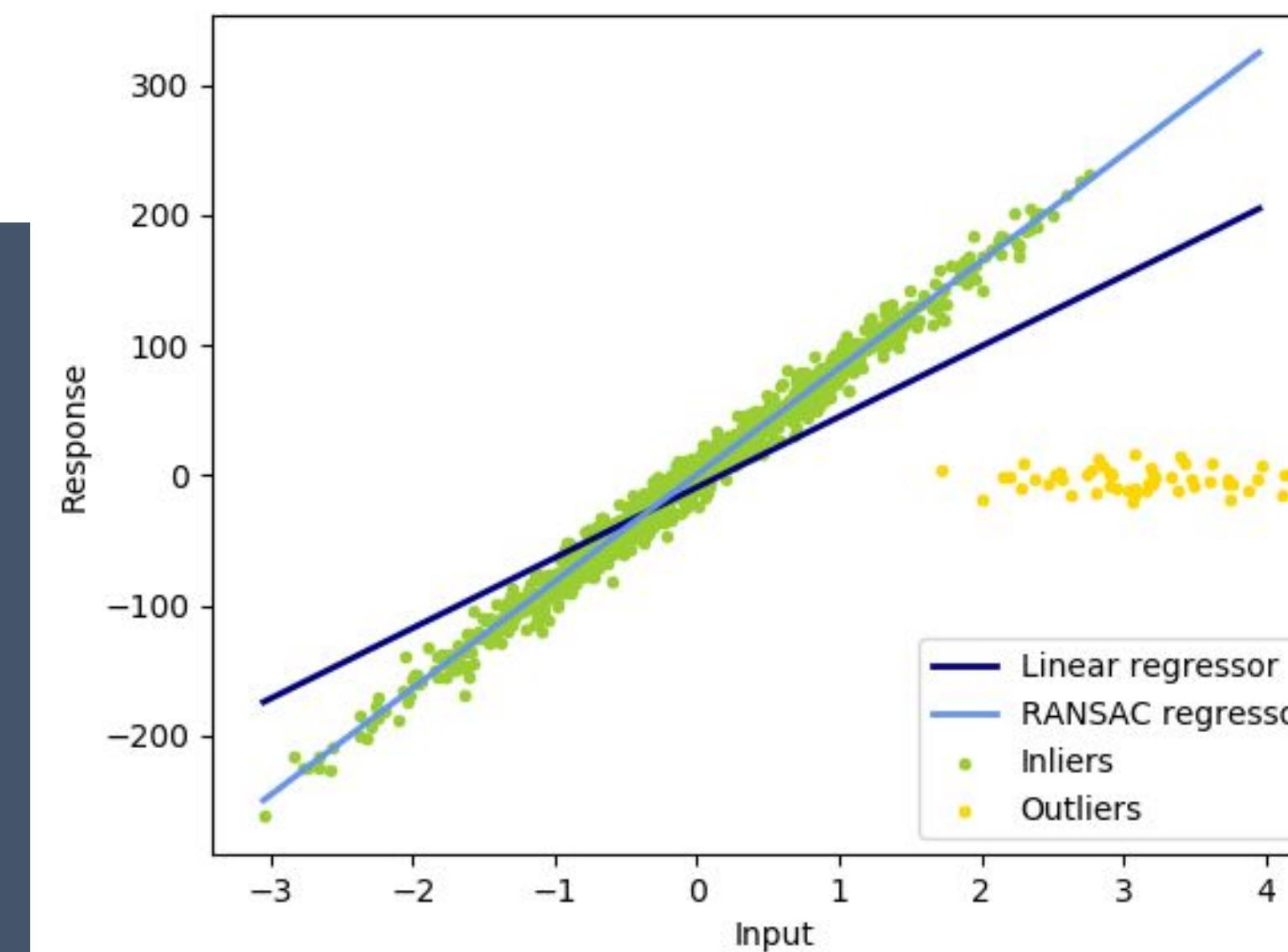
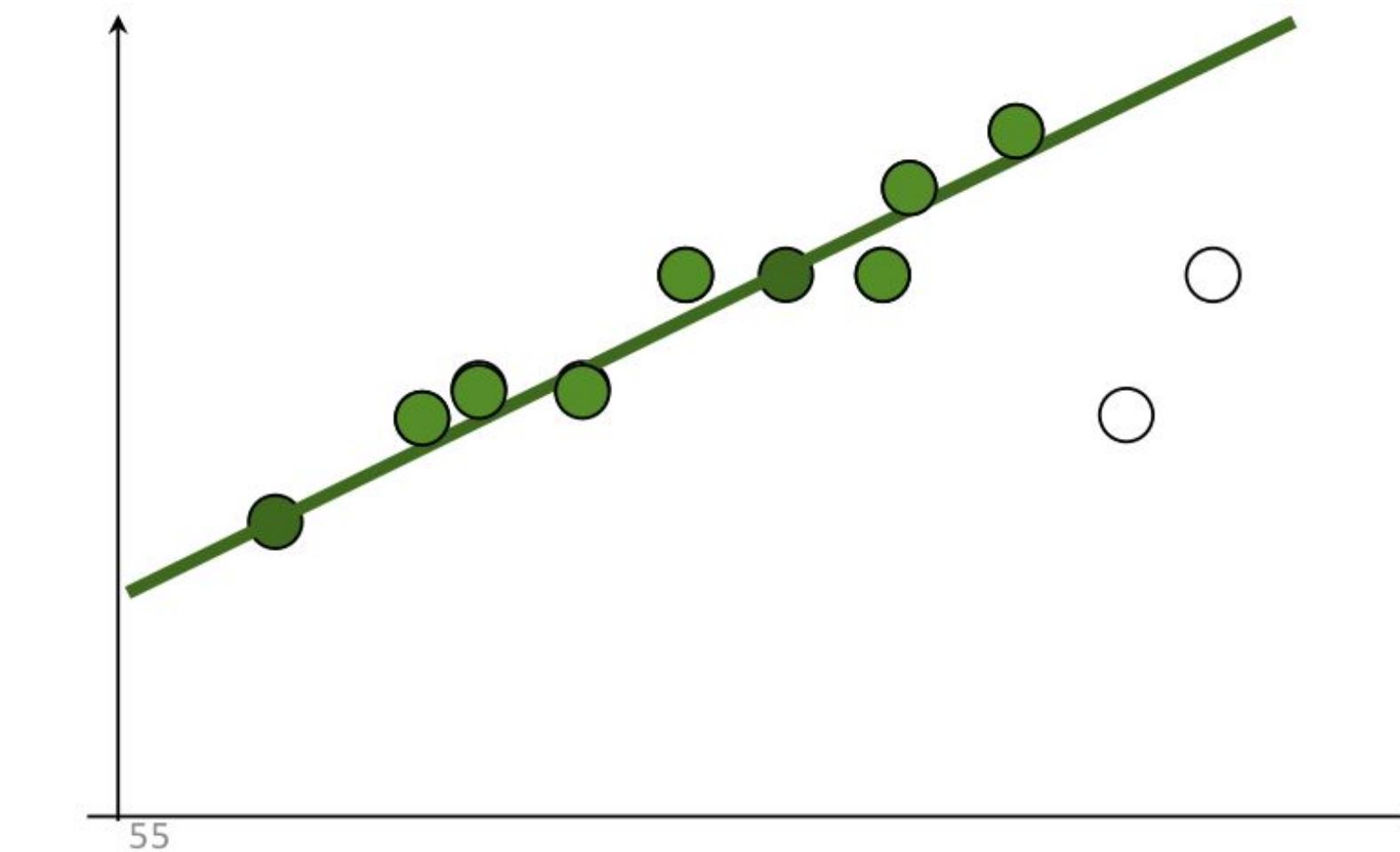


# Remapeamento de imagem (Image Warping)

- Escolha 2 pontos
- Ajuste a linha
- Conte inliers



- Use o maior conjunto de inliers
- Ajuste o mínimo quadrado



# Remapeamento de imagem (Image Warping) e Blending

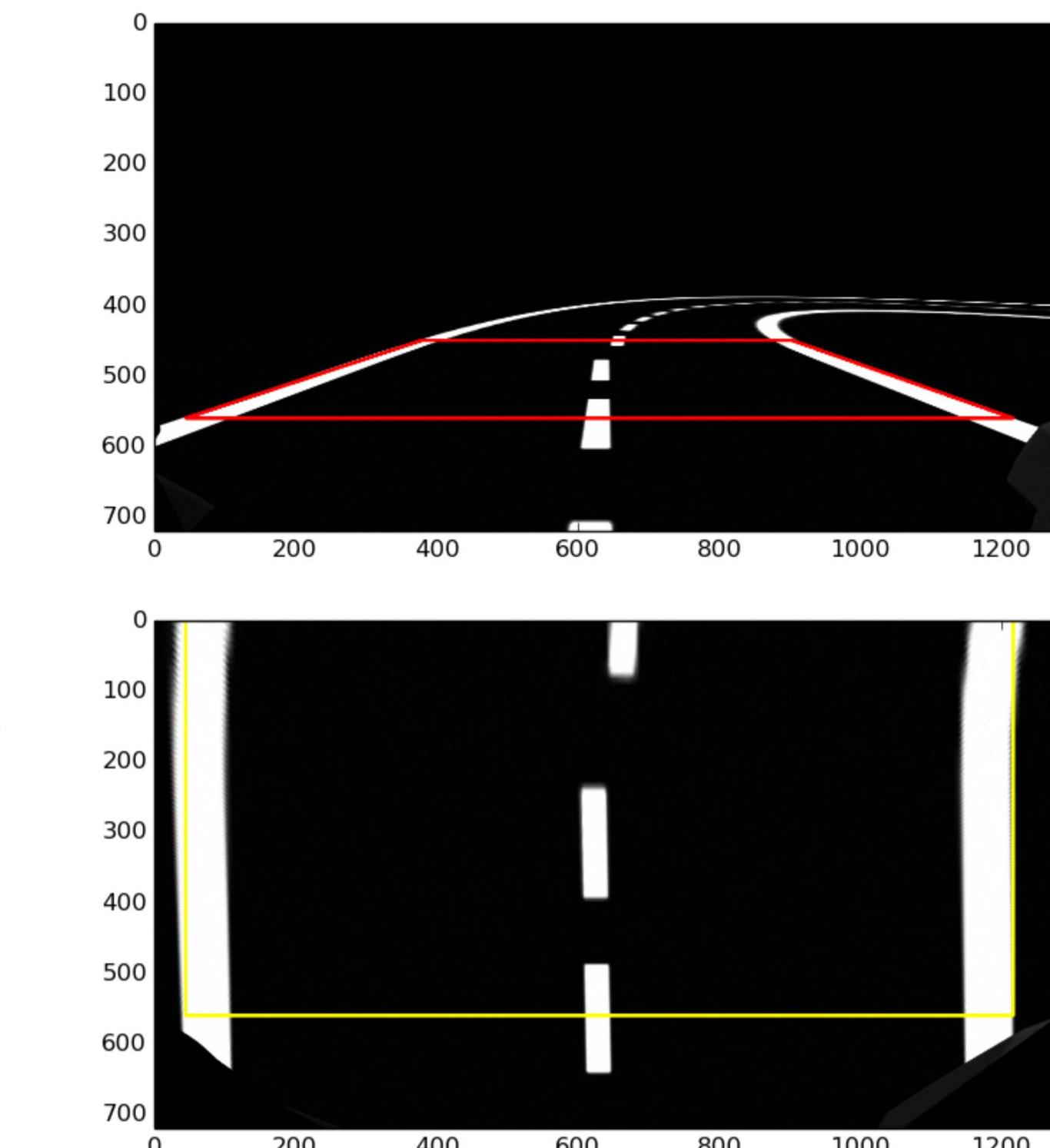
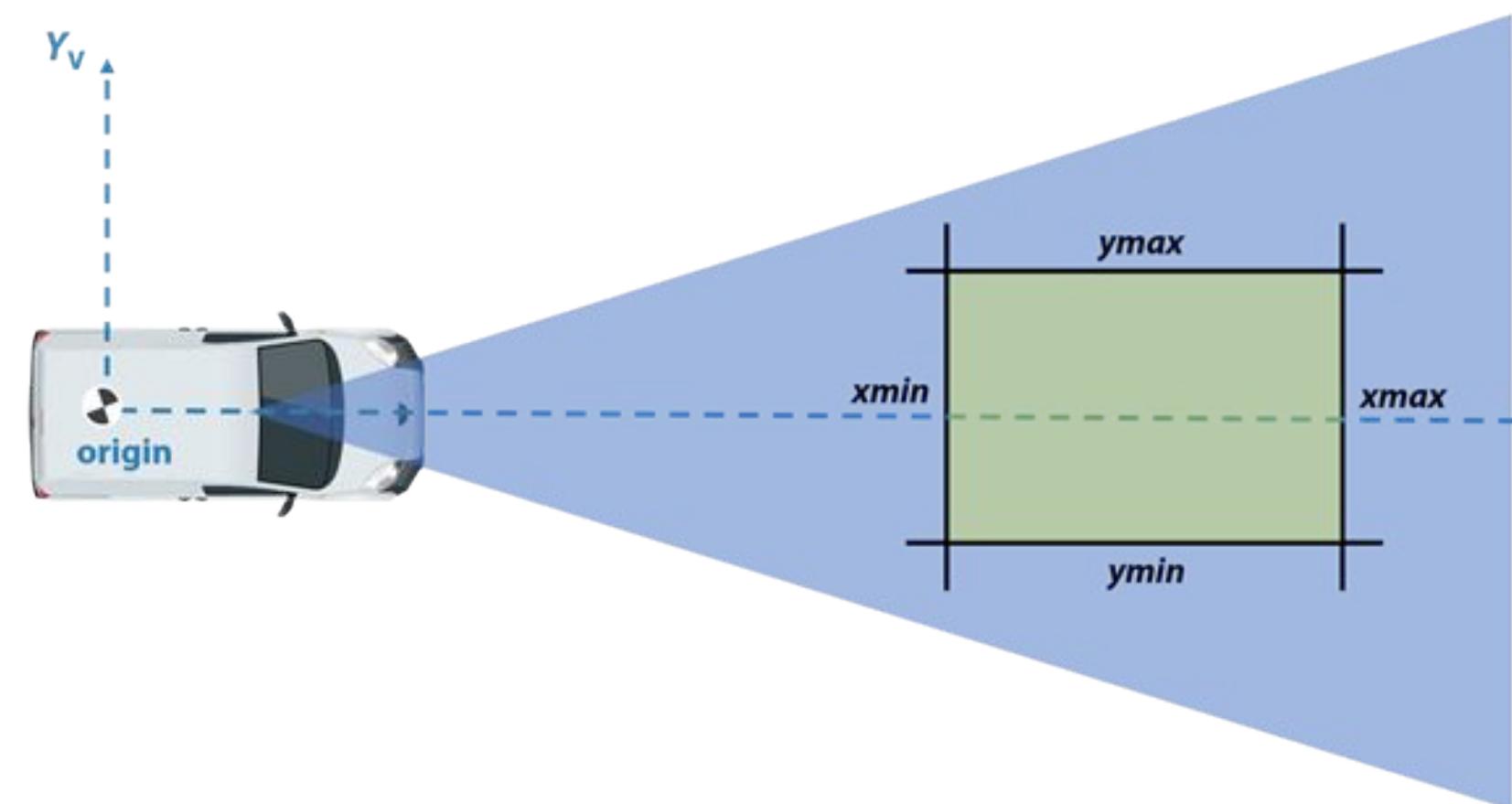
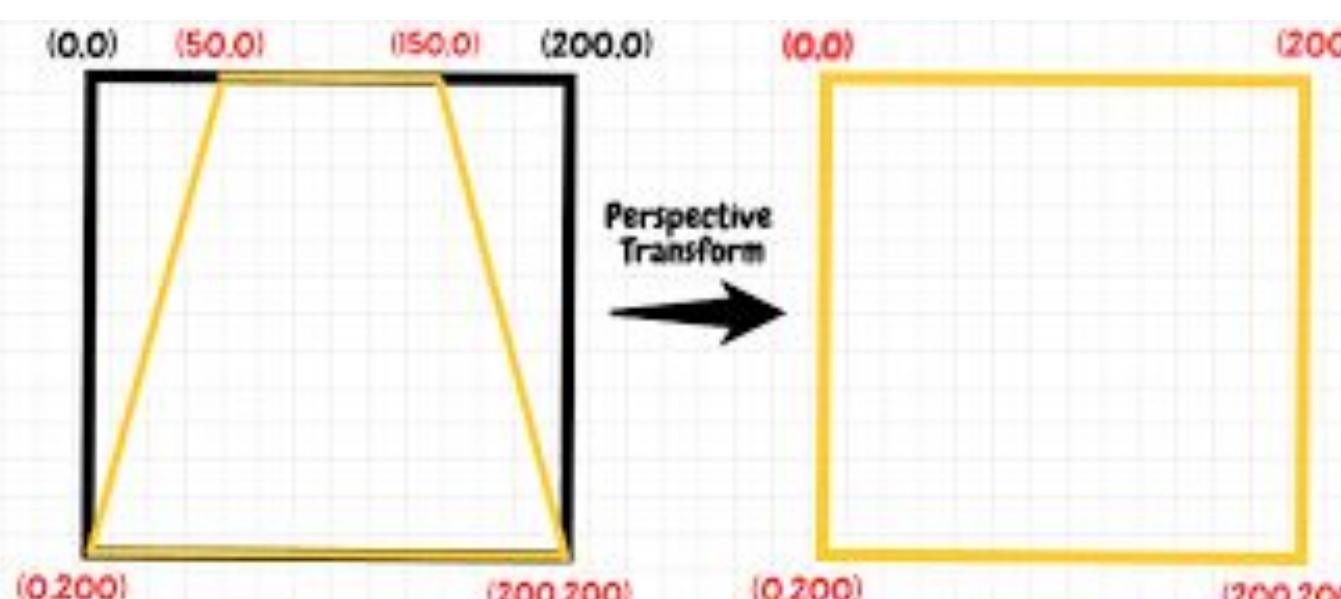
- Detecção de bordas da imagem inicial
- Transformação de perspectiva (Perspective Transform)
- Concatena bordas da imagem inicial e seguinte
- Executa uma translação em H
- Aplica o Warping

$$\begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

$$x' = x + t_x$$

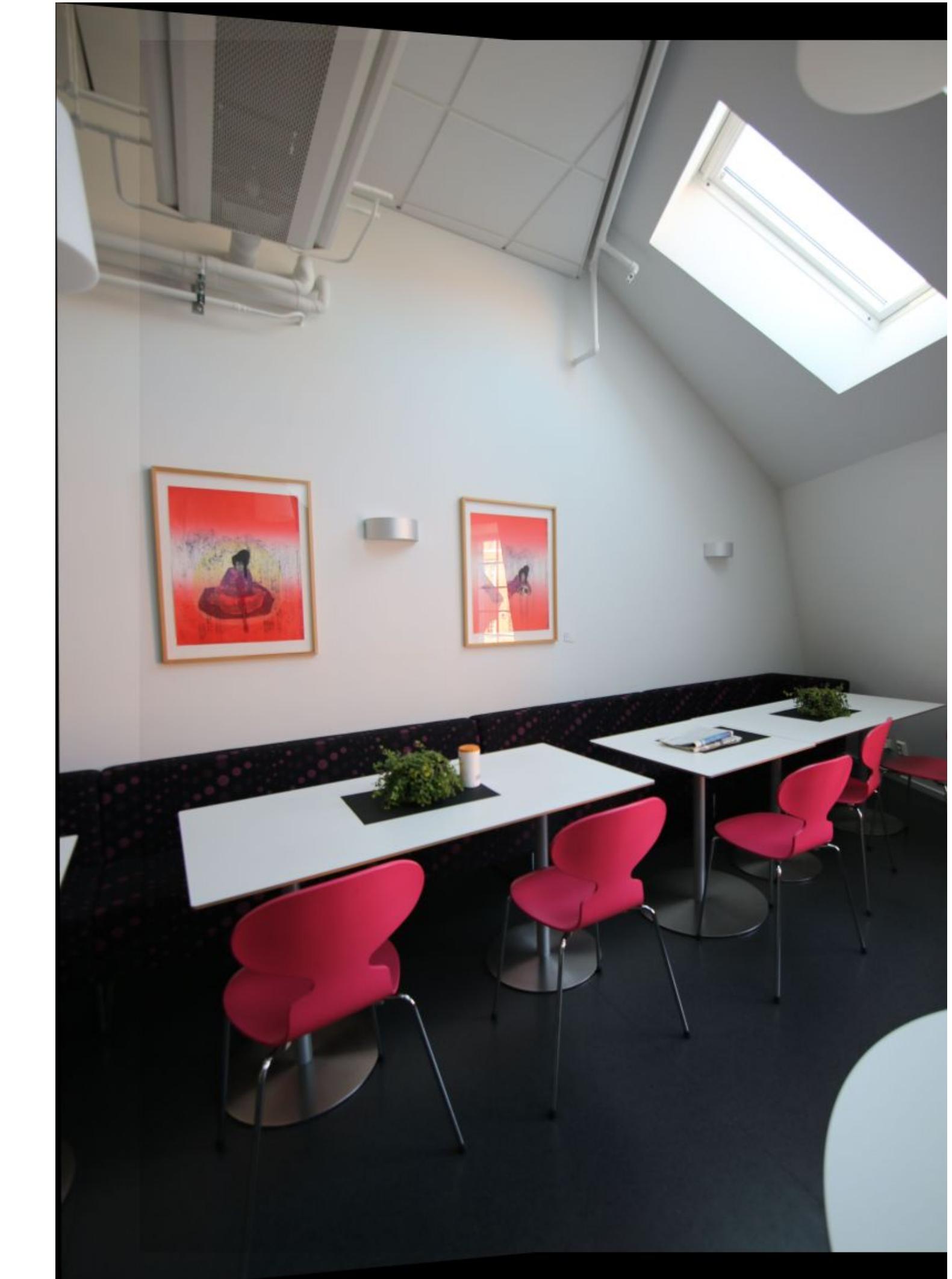
$$y' = y + t_y$$

- Translação
- $t_x = -\text{xmin}(\text{matriz bordas})$
- $t_y = -\text{ymin}(\text{matriz bordas})$



## Remapeamento de imagem (Image Warping) e Blending

- $t_x = 76$   
 $t_y = 34$
- ```
result = warp(img)
new_img -> result[t_y:altura_proxima_imagem+t_y, t_x:altura_proxima_imagem+t_x] = proxima_imagem
```



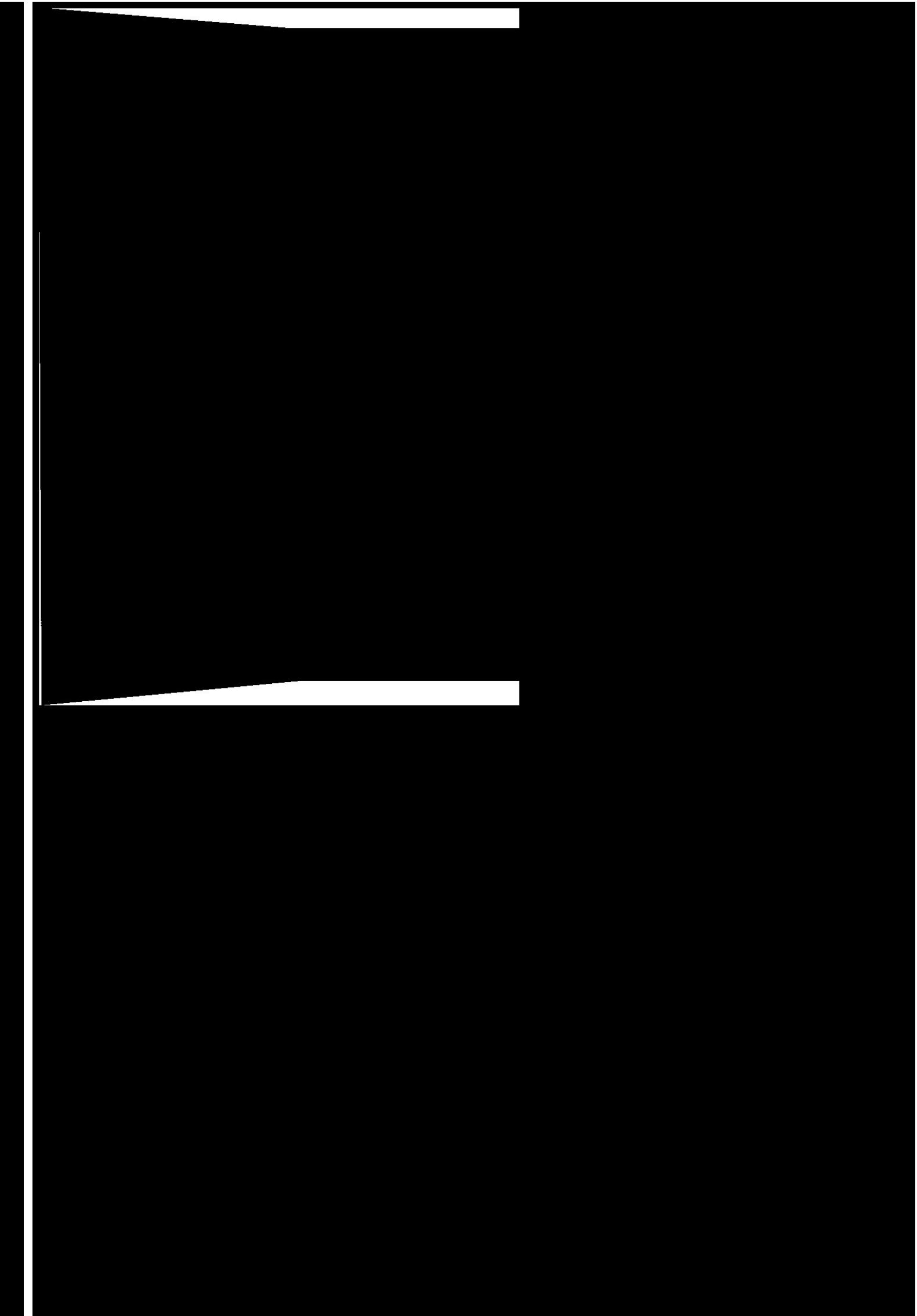
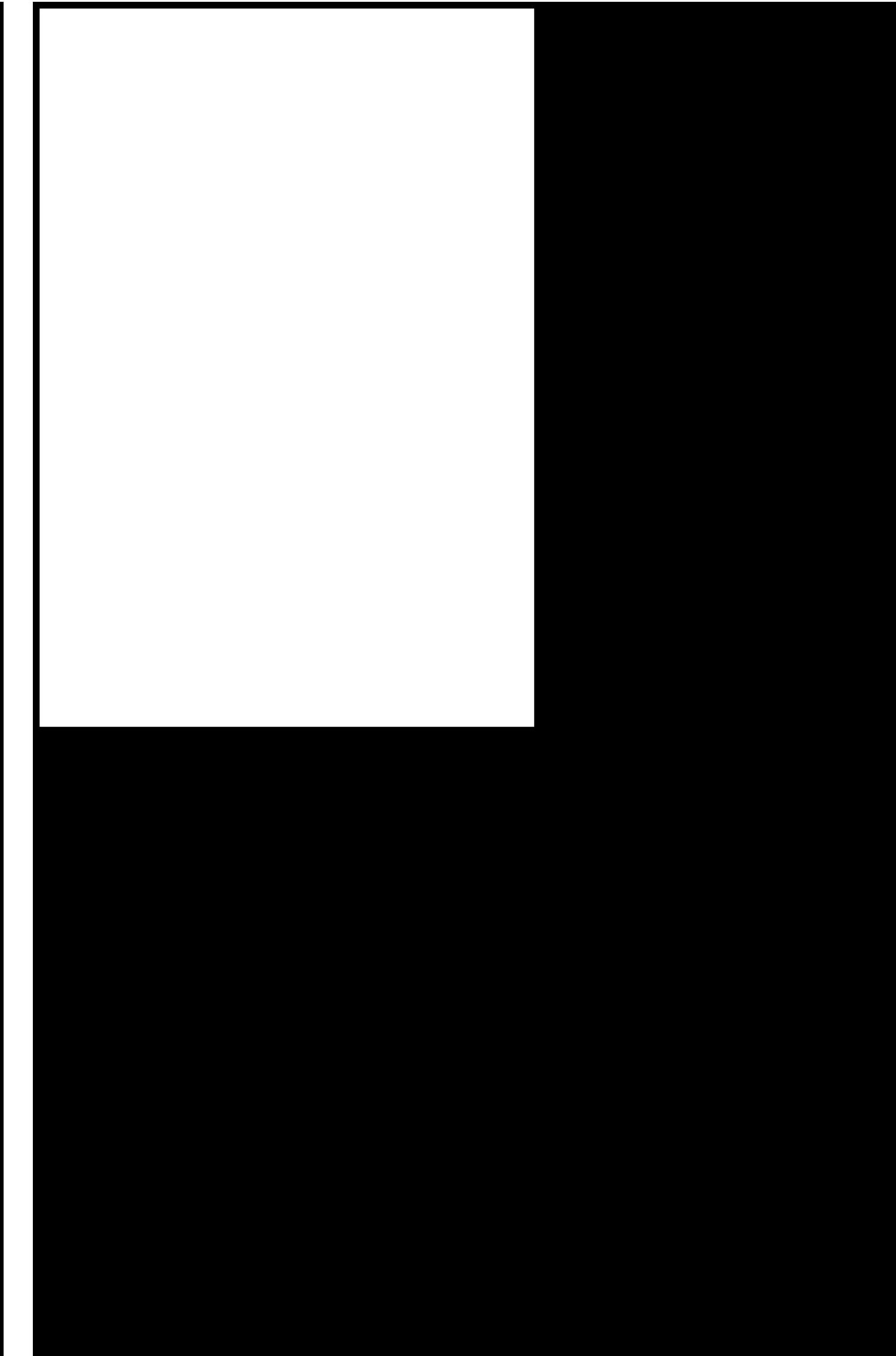
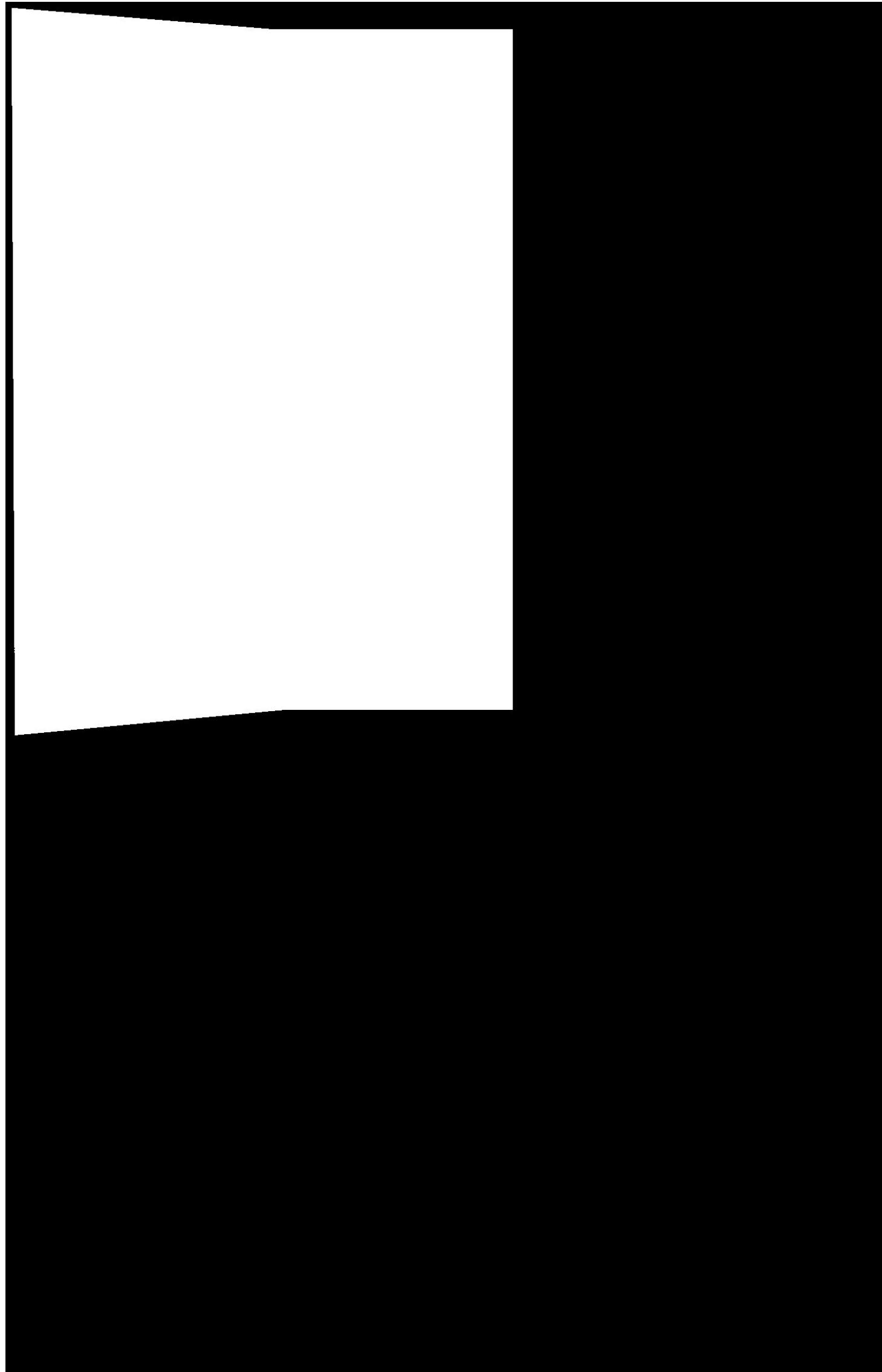
# Remapeamento de imagem (Image Warping) e Blending

- Cropping



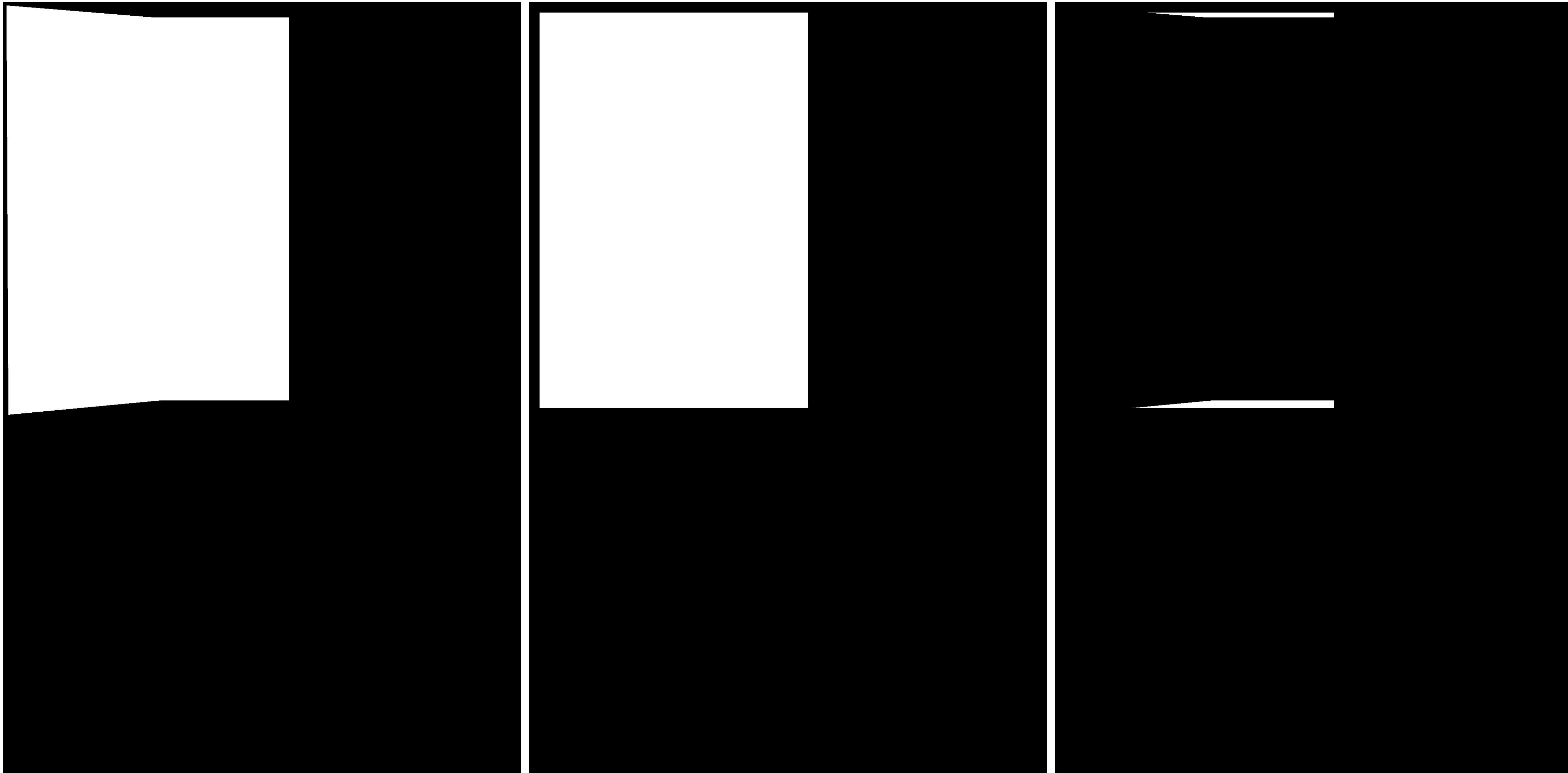
# Remapeamento de imagem (Image Warping) e Blending

- Cropping -- Passo 1



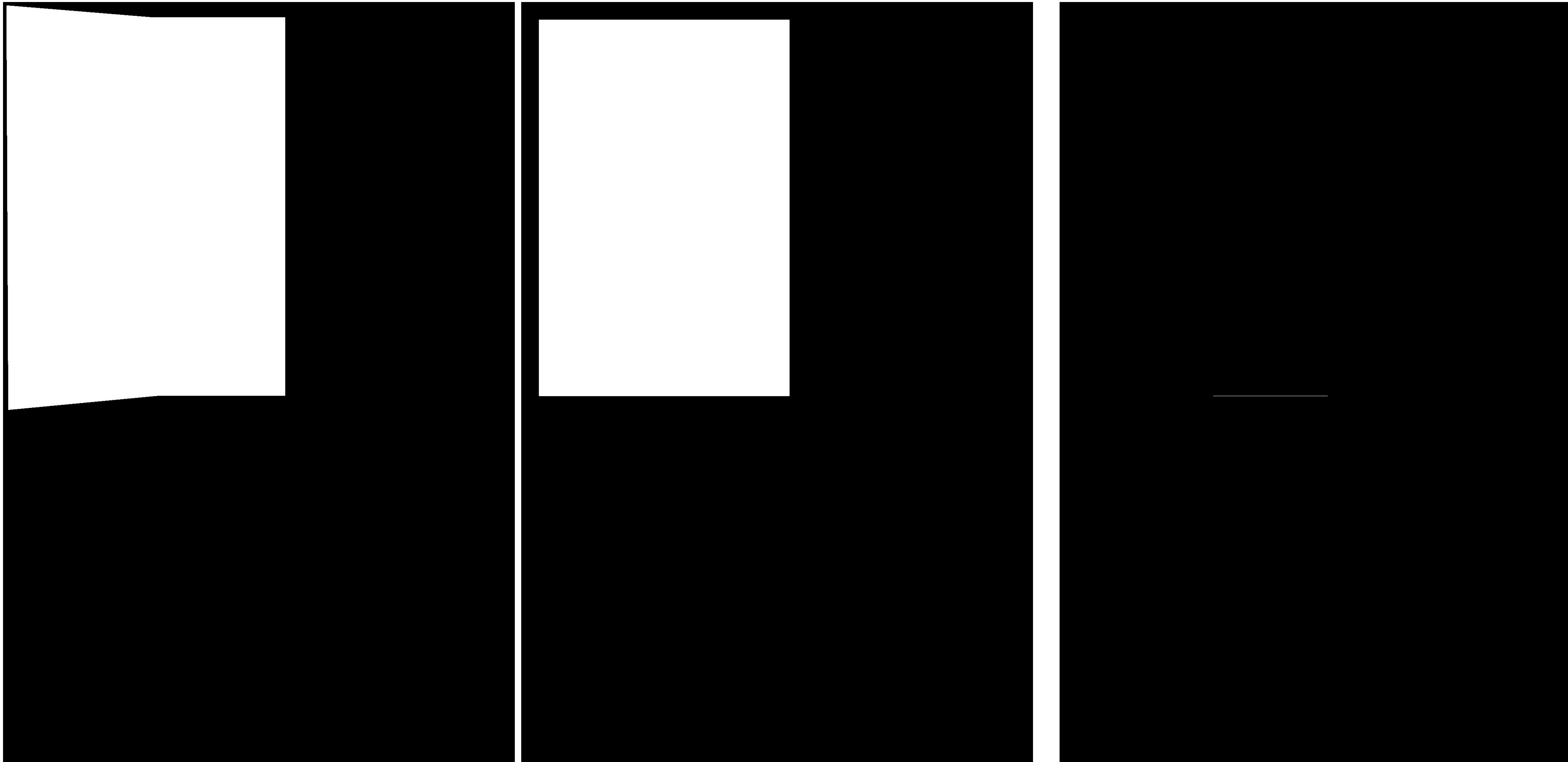
# Remapeamento de imagem (Image Warping) e Blending

- Cropping -- Passo 20



# Remapeamento de imagem (Image Warping) e Blending

- Cropping -- Passo 41



# Remapeamento de imagem (Image Warping) e Blending

- Resultado

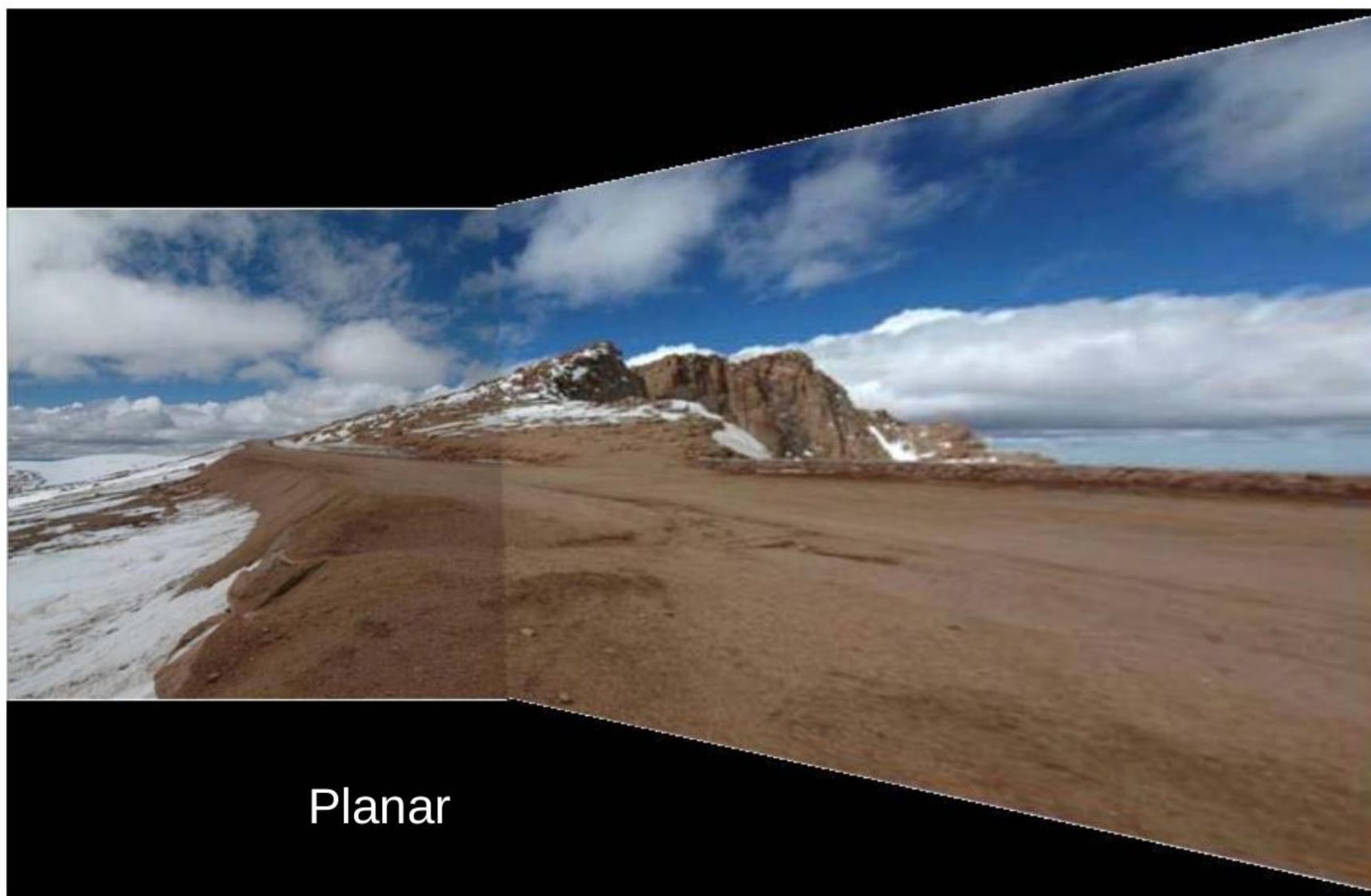


- Resultado Opencv



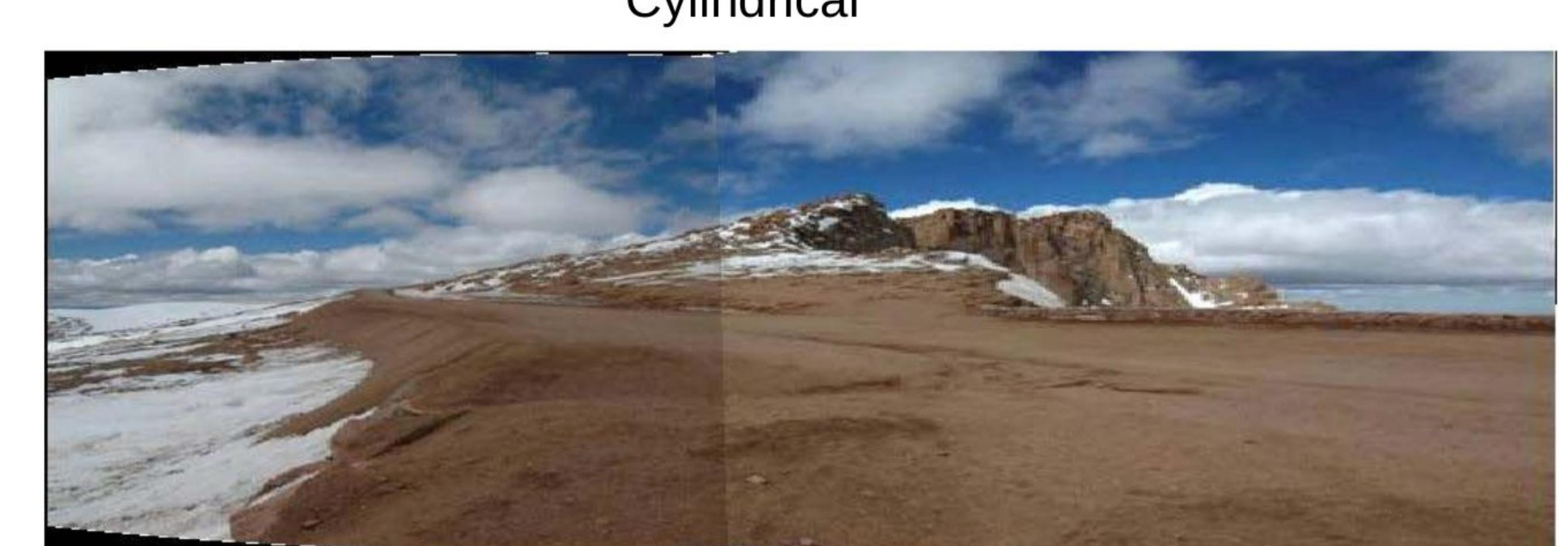
# Remapeamento de imagem (Image Warping) e Blending

- Resultado



Planar

- Resultado Opencv



Cylindrical

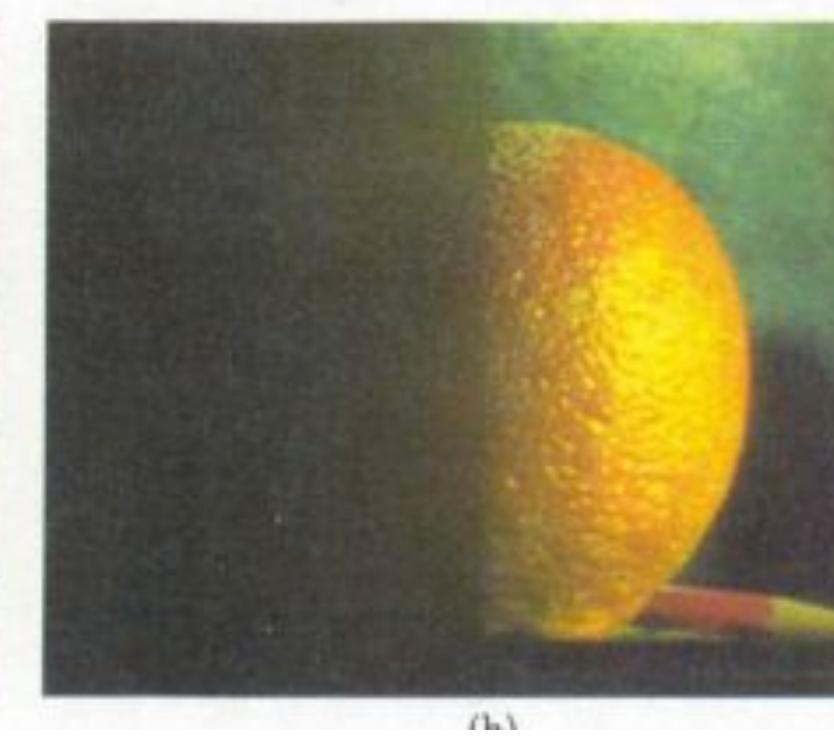
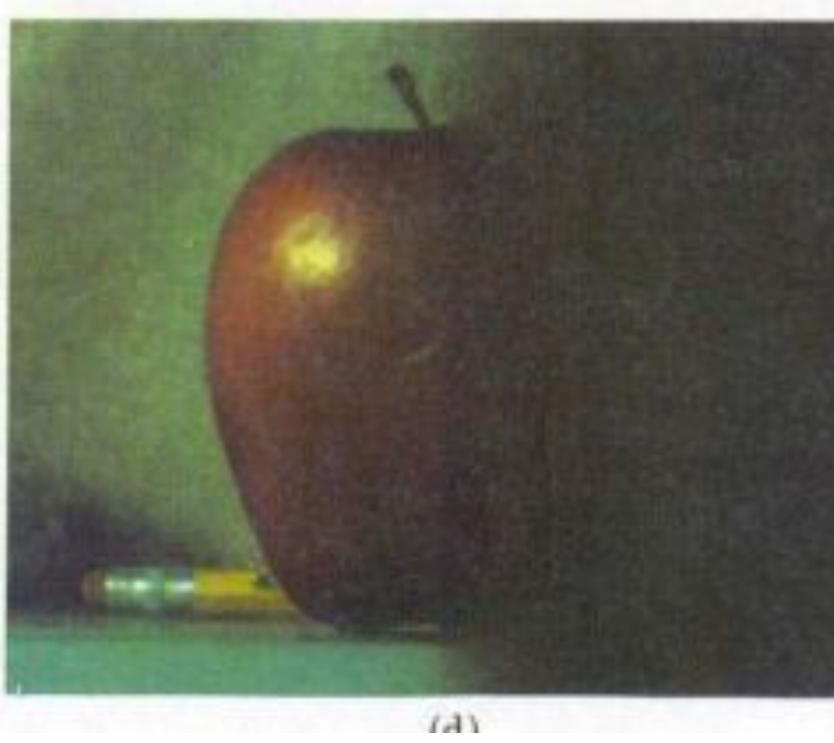
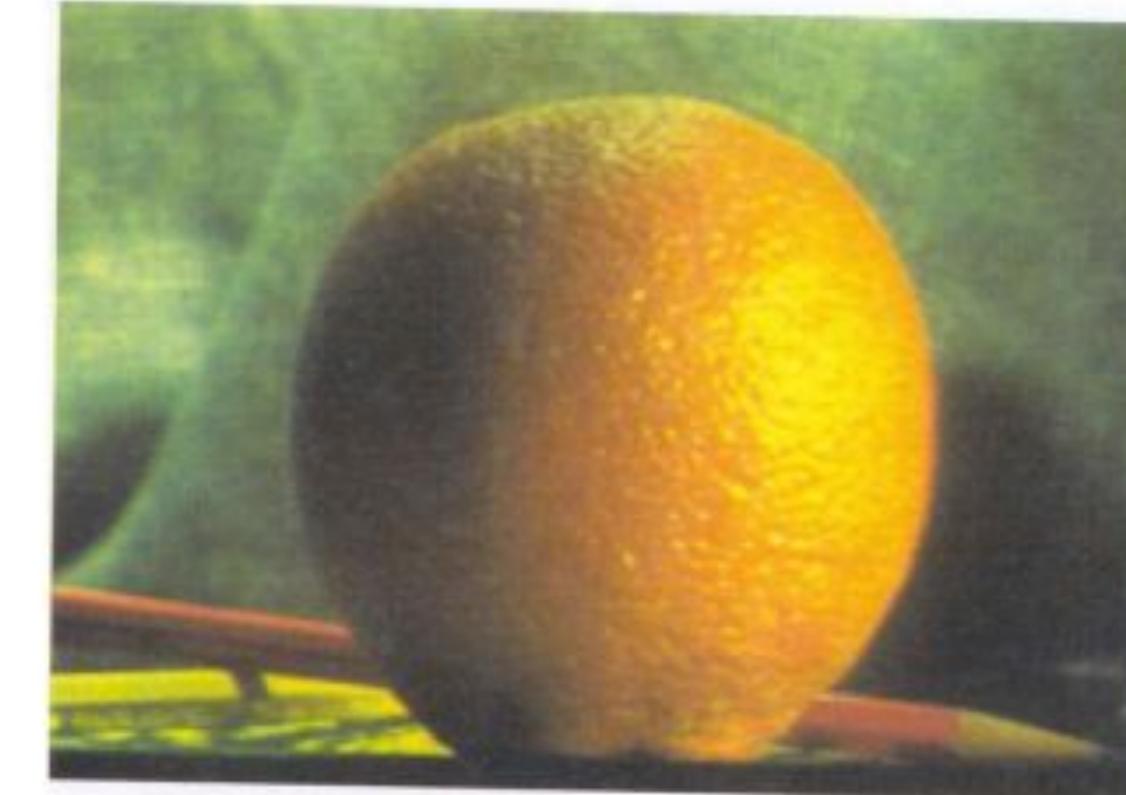
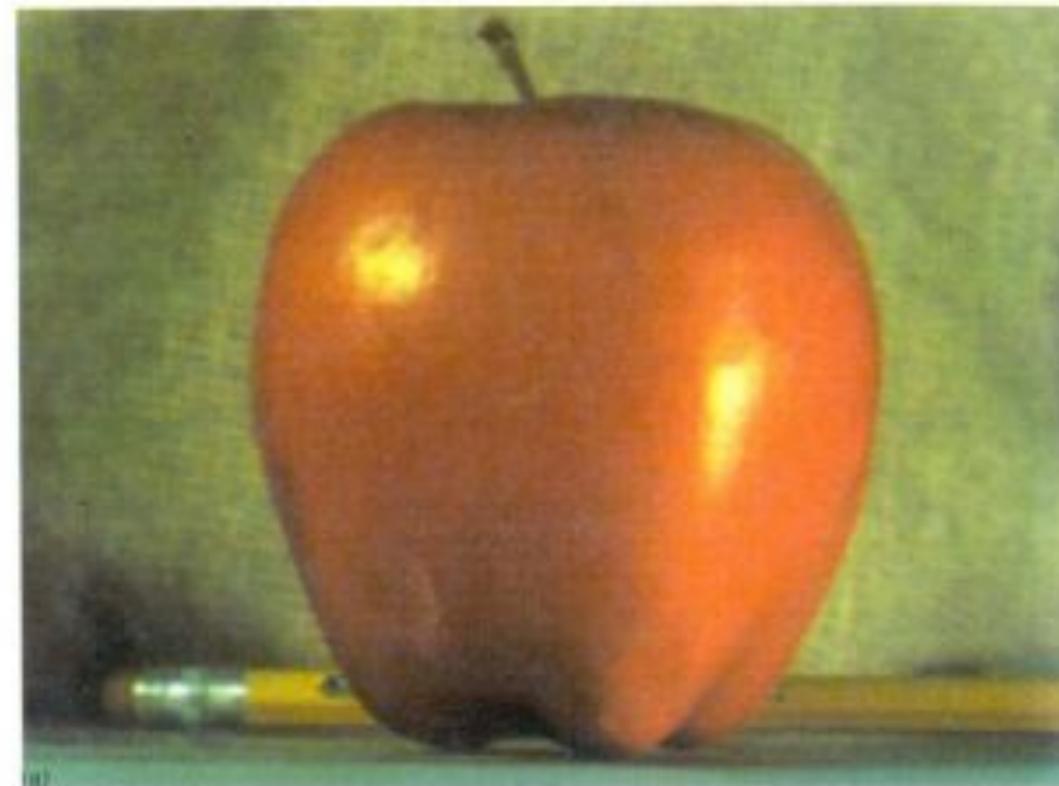
# Remapeamento de imagem (Image Warping) e Blending

- Alfa blending



# Remapeamento de imagem (Image Warping) e Blending

- Pyramid blending



# Remapeamento de imagem (Image Warping) e Blending

- Two-Band blending



# Remapeamento de imagem (Image Warping) e Blending

---

- Desafios

- Descoberta e combinação de keypoints e features

- Desafios

- Supressão de paralaxe: Ruído como um aparente deslocamento em um objeto observado na imagem.
- Suavização de intersecções



- Métodos baseados em blocos de características.
- Métodos para correção de cores em imagem com paralaxe.
- Métodos baseados em segmentação superpixel

# Referências

---

- <https://stackabuse.com/affine-image-transformations-in-python-with-numpy-pillow-and-opencv/>
- <https://www.learnopencv.com/image-alignment-feature-based-using-opencv-c-python/>
- Wang, C., Gao, Z., & Lu, Q. (2020, July). Parallax-Based Color Correction in Image Stitching. In *2020 IEEE 5th International Conference on Image, Vision and Computing (ICIVC)* (pp. 69-74). IEEE.
- Shi, H., Guo, L., Tan, S., Li, G., & Sun, J. (2019). Improved parallax image stitching algorithm based on feature block. *Symmetry*, 11(3), 348.
- Zhao, C., Zhang, H., Chen, J., & Fu, W. (2019, May). Region-based parallax-tolerant image stitching. In *Tenth International Conference on Graphics and Image Processing (ICGIP 2018)* (Vol. 11069, p. 1106909). International Society for Optics and Photonics.
- Yang, X., Liu, Z. Y., Qiao, H., Su, J. H., Ji, D. X., Zang, A. Y., & Huang, H. (2020). Graph-Based Registration and Blending for Undersea Image Stitching. *Robotica*, 38(3), 396-409.
- Zhang, Y., Lai, Y. K., & Zhang, F. L. (2019). Stereoscopic image stitching with rectangular boundaries. *The Visual Computer*, 35(6-8), 823-835.
- Saha, A., Maity, S., & Bhowmick, B. (2019, May). Multi-modal Image Stitching with Nonlinear Optimization. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1987-1991). IEEE.
- Muja, M., & Lowe, D. G. (2014). Scalable nearest neighbor algorithms for high dimensional data. *IEEE transactions on pattern analysis and machine intelligence*, 36(11), 2227-2240.
- Liu, Y., Zhang, H., Guo, H., & Xiong, N. N. (2018). A fast-brisk feature detector with depth information. *Sensors*, 18(11), 3908.
- Leutenegger, S., Chli, M., & Siegwart, R. Y. (2011, November). BRISK: Binary robust invariant scalable keypoints. In *2011 International conference on computer vision* (pp. 2548-2555). Ieee.
- Lindeberg, T. (2012). Scale invariant feature transform.
- Brown, M., & Lowe, D. G. (2007). Automatic panoramic image stitching using invariant features. *International journal of computer vision*, 74(1), 59-73.
- <https://courses.cs.washington.edu/courses/cse576/17sp/notes/>
- Mehta, J. D., & Bhirud, S. G. (2011). Image stitching techniques. In *Thinkquest~ 2010* (pp. 74-80). Springer, New Delhi.
- Joshi, K. A Survey on Real Time Image Stitching.
- Brown M. (2020) Image Stitching. In: Ikeuchi K. (eds) Computer Vision. Springer, Cham. [https://doi.org/10.1007/978-3-030-03243-2\\_13-1](https://doi.org/10.1007/978-3-030-03243-2_13-1)
- <https://courses.engr.illinois.edu/cs498dwh/fa2010/>
-

---

Obrigado

Marcos Oliveira

[github.com/MarcosSoares10](https://github.com/MarcosSoares10)  
[www.linkedin.com/in/marcos-soares-391718141](https://www.linkedin.com/in/marcos-soares-391718141)