
Lecture #5: Panoramização, modelos de movimento, alinhamento de imagens e composição

Alexandre Yuji Kajiahira, Noemi Pereira Scherer
Department of Computer Science
Federal University of Technology - Paraná / UTFPR
Campo Mourão, Paraná, Brazil
{alexandre.ykz, noemischerer13}@gmail.com

1 Introdução

Imagens panorâmicas são imagens de alta resolução, *stitching* com múltiplas imagens. Para criar esse tipo de imagem é necessário utilizar modelos de movimentos (que é através de um modelo simples de transformações 2D ou por meio de um modelo de perspectiva planar, etc) dependendo do tipo de *stitching*. Após aplicar um dos modelos de movimento, recebemos um par de imagens para registrar, com o objetivo de encontrar um conjunto de parâmetros de alinhamento que minimize o erro de registro entre todos os pares de imagens. Por fim, com a composição iremos construir a imagem final, que envolve a escolha da superfície de composição final, a composição da vista e também envolve a maneira de como misturar de forma ideal esses *pixels* para minimizar costuras visíveis, borrões e "fantasmas". A seguir, nas seções abaixo iremos nos aprofundar mais fundo na panoramização, modelos de movimentos, alinhamento e composição.

2 Panoramização

Imagens panorâmicas são fotografias de alta resolução ¹, em que é realizado o *stitching* a partir de múltiplas fotos ². As imagens panorâmicas aumentam efetivamente o campo de visão de uma câmera[5], capturam cenas que não podem ser capturadas com um *frame*, retratam cenas amplas [11], é mais barato e é mais fácil de alcançar efeitos que de alto custo. Se desejamos criar uma imagem panorâmica a partir de duas fotos sobrepostas, precisamos mapear um plano de imagem para o outro e, para ser panoramas de boa qualidade, a transformação entre as imagens precisa ser mais possível [11]. Assim, como o HDR (*High Dynamic Range*) e o *Focus Stacking* também usam métodos computacionais para ir além das limitações físicas da câmera [3].

3 Modelos de movimento

Antes de registrar e alinhar as imagens, é necessário estabelecer, as relações matemáticas que mapeiam as coordenadas dos *pixels* de uma imagem para outra [9]. Os modelos de movimentos são uma representação numérica dos movimentos da câmera entre duas imagens [4].

¹Uma imagem panorâmica de alta qualidade depende do número de imagens de entradas [2]

²A ideia do *stitching* das imagens é combinar duas ou mais imagens para gerar uma maior [3], porém alguns problemas com imagens panorâmicas são, as distorções de lentes, os movimentos de cena e as diferentes exposições [6].

Nos algoritmos de *stitching*, é utilizado as homografias, um ponto $x = (u, v, 1)$ na imagem um e $x' = (u', v', 1)$ em outra imagem, a homografia é uma matriz três por três M .

$$M = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} \quad (1)$$

Essa matriz de homografia relaciona os *pixels* das coordenadas nas duas imagens se $x' = x$. Quando aplicada a cada pixel, a nova imagem é a versão distorcida da imagem original [7]. Então, nos algoritmos de *stitching*, as homografias são usadas para obter os modelos de movimentos [4]. Além disso, esses modelos podem ser aplicados a diferentes situações de *stitching* [9], por exemplo, o *stitching* em um vídeo precisa de dois tipos de modelos de movimento, que são o modelo de movimento espacial³ e temporal⁴, enquanto uma imagem precisa apenas do modelo de movimento espacial [4]. Existem modelos de movimentos paramétricos, desde simples transformações 2D, modelo em perspectiva planares, rotações de câmeras 3D, distorções de lentes e mapeamento para superfícies não planas [9], etc. Nas subseções abaixo, iremos mostrar alguns desses dois modelos.

3.1 Movimento de perspectiva planar

Esse tipo de modelo de movimento é o mais simples a ser usado para alinhar as imagens em que simplesmente translaciona e rotaciona em 2D, e é o mesmo utilizado se tivesse impressões fotográficas sobrepostas. A translação e a rotação podem ser usadas em modelos de movimento adequados para compensar pequenos movimentos da câmera em aplicações como estabilização de fotos, vídeos e mesclagem [9].

Nesse tipo de modelo de movimento, o mapeamento entre duas câmeras visualizando um plano comum é descrito usando uma homografia. Essa técnica formou a base de alguns dos algoritmos automatizados, como as técnicas confiáveis de correspondência de características ainda não haviam sido desenvolvidas, esses algoritmos usaram a correspondência direta de valores de *pixel*, isto é, a estimativa de movimento paramétrico direto. Algoritmos de costura mais recentes extraem recursos e os combinam, geralmente usando técnicas robustas como o RANSAC (*RANdom SAmple Consensus*) para calcular um bom conjunto de *inliers* [9].

3.2 Panoramas rotacionais

O caso mais comum de imagem de *stitching* é quando a câmera passa por uma rotação pura, que acontece em relação a geometria distante da cena, à medida que você se afasta, a câmera passa por essa rotação, o que equivale a assumir que todos os pontos estão muito longe da câmera, ou seja, no plano infinito. Para isso, definimos $t_0 = t_1 = 0$, obtemos a matriz de homografia simplificada 3x3

$$\tilde{H}_{10} = K_1 R_1 R_0^{-1} K_0^{-1} = K_1 R_{10} K_0^{-1} \quad (2)$$

onde $K_k = \text{diag}(f_k, f_k, 1)$ é a matriz intrínseca de câmera simplificada, assumindo $c_x = c_y = 0$, ou seja, estamos indexando os *pixels* a partir do centro óptico, que é reescrito como

$$\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \sim \begin{bmatrix} f_1 & & \\ & f_1 & \\ & & 1 \end{bmatrix} R_{10} \begin{bmatrix} f_0^{-1} & & \\ & f_0^{-1} & \\ & & 1 \end{bmatrix} \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} \quad (3)$$

ou

$$\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \sim R_{10} \begin{bmatrix} x_0 \\ y_0 \\ 1 \end{bmatrix} \quad (4)$$

³São homografias entre os movimentos das câmeras a qual representa homografias entre múltiplas entradas de um mesmo *frame* [4].

⁴Consiste entre sucessivos movimentos da câmera e do caminho da câmera, onde o movimento da câmera são homografias entre os vizinhos de *frames* da mesma câmera, e o caminho da câmera são os componentes importantes para acurácia de costura do vídeo [4].

que revela a simplicidade das equações de mapeamento e torna todos os parâmetros de movimentos explícitos. Assim, em vez da homografia geral de oito parâmetros que relaciona um par de imagens, obtemos modelos de movimento de rotação 3D de três, quatro ou cinco parâmetros correspondentes aos casos em que a distância focal f é conhecida, fixa ou variável. Estimar a matriz de rotação 3D associada a cada imagem é intrinsecamente mais estável do que estimar uma homografia com oito graus de liberdade, o que faz dele o método de escolha para algoritmos de costura de imagem em larga escala. Dada uma estimativa atual para a homografia \tilde{H}_{10} (na equação 3), que é a melhor maneira de atualizar R_{10} é pré-adicionar uma matriz de rotação incremental $R(\omega)$ para estimar atualmente R_{10}

$$\tilde{H}(\omega) = K_1 R(\omega) R_{10} K_0^{-1} = [K_1 R(\omega) K_1^{-1}] [K_1 R_{10} K_0^{-1}] = D \tilde{H}_{10} \quad (5)$$

Note que escrevemos a regra de atualização na forma de composição, onde a atualização incremental D é prefixada à homografia atual \tilde{H}_{10} . Usando a aproximação de pequeno ângulo para $R(\omega)$, então a matriz de atualização incremental pode escrita como,

$$D = K_1 R(\omega) K_1^{-1} \approx K_1 (I + [\omega]_x) K_1^{-1} = \begin{bmatrix} 1 & -\omega_z & f_1 \omega_y \\ \omega_z & 1 & -f_1 \omega_x \\ -\omega_y/f_1 & \omega_x/f_1 & 1 \end{bmatrix} \quad (6)$$

Observe que existe uma correspondência um-para-um entre as entradas na matriz D e o h_{00}, \dots, h_{21} .

$$(h_{00}, h_{01}, h_{02}, h_{10}, h_{11}, h_{12}, h_{20}, h_{21}) = (0, -\omega_z, f_1 \omega_y, \omega_z, 0, -f_1 \omega_x, -\omega_y/f_1, \omega_x/f_1) \quad (7)$$

Então combinando a equação acima, com a equação 8 abaixo, teremos a seguinte equação 9.

$$\begin{bmatrix} \hat{x}' & -x \\ \hat{y}' & -y \end{bmatrix} = \begin{bmatrix} x & y & 1 & 0 & 0 & 0 & -x^2 & -xy \\ 0 & 0 & 0 & x & y & 1 & -xy & y^2 \end{bmatrix} \begin{bmatrix} \Delta h_{00} \\ \vdots \\ \Delta h_{21} \end{bmatrix} \quad (8)$$

$$\begin{bmatrix} \hat{x}' & -x \\ \hat{y}' & -y \end{bmatrix} = \begin{bmatrix} -xy/f_1 & f_1 + x^2/f_1 & -y \\ -(f_1 + y^2/f_1) & xy/f_1 & x \end{bmatrix} \begin{bmatrix} \omega_x \\ \omega_y \\ \omega_z \end{bmatrix} \quad (9)$$

A equação 9 fornece atualizações linearizadas necessárias para estimar $\omega = (\omega_x, \omega_y, \omega_z)$, esta regra de atualização depende da distância focal f_1 da vista de destino e independe da distância focal f_0 da vista de modelo. Isso acontece porque o algoritmo de composição faz essencialmente pequenas perturbações no alvo. Uma vez o vetor de rotação incremental foi calculado, a matriz de rotação R_1 é realizada $R_1 \leftarrow R(\omega) R_1$ [9].

3.3 Coordenadas cilíndricas e esféricas

Uma alternativa ao uso de homografias ou movimentos 3D para alinhar imagens é primeiro transformar as imagens em coordenadas cilíndricas e depois usar um modelo puramente translacional para alinhá-las. Infelizmente os algoritmos de costura cilíndrica só funciona se as imagens forem tiradas com uma câmera nivelada e girar apenas em torno do seu eixo vertical ou com ângulo de inclinação conhecido. Sob essas condições, imagens em rotações diferentes são relacionadas por uma translação horizontal pura [9].

3.4 Gap closing

O *Gap closing* é usado para estimar uma série de matrizes de rotação e distâncias focais, que são encadeadas para criar grandes panoramas. Porém, essa abordagem não produzirá um panorama fechado de 360°, em vez disso haverá uma lacuna ou sobreposição. Uma maneira de resolver esse problema seria combinar a primeira imagem na sequência com a última. A diferença entre as duas estimativas de matriz de rotação associadas à primeira repetida indica a quantidade de erros de registro. Porém, existirá uma lacuna entre as imagens que pode ser usada uma distância focal, que só funciona para um panorama "unidimensional" em uma câmera que está girando continuamente na mesma direção [9].

3.5 Distorções de lentes

Ao tirar imagens com lentes grande angular, muitas vezes é necessário modelar as distorções de lentes, como a distorção radial. Essa distorção diz que as coordenadas nas imagens observadas são deslocadas (distorção de barril) ou para (distorção de almofada) o centro da imagem por uma quantidade proporcional à sua distância radial. Esse tipo de modelo mais simples usam polinômios de baixa ordem [8].

Várias técnicas podem ser usadas para estimarem os parâmetros de distorção radial para uma determinada lente. O mais simples e útil é tirar uma imagem de uma cena com muitas linhas retas, especialmente linhas alinhadas e próximas às bordas da imagem, assim os parâmetros podem então ser ajustados até que todas as linhas de imagem fiquem retas, esse processo é chamado de *plumb*. Uma outra alternativa é usar várias imagens sobrepostas e combinar a estimativa dos parâmetros de distorção radial junto com o processo de alinhamento de imagem. No caso da lente *fish eye* (olho de peixe) requer um modelo diferente dos modelos polinomiais tradicionais de distorção radial. Essas lentes se comportam, em uma primeira aproximação, como projetores de distância equidistantes de ângulos distantes do eixo óptico, que é o mesmo que a projeção polar [8].

3.6 Aplicações

Algumas aplicações que utilizam esses modelos são: quadro branco, *scanner* de documentos, sumarização de vídeo e compactação de vídeo [9].

Um exemplo utilizando o quadro branco, é quando você tira várias fotos da mesa sobrepostas de um quadro branco ou outro objeto grande plano [9].

Já no *scanner* de documentos, o aplicativo de *stitching* de imagem mais simples une várias digitalizações de imagem feitas em uma *scanner* mesa. Devem ser feitas várias digitalizações do documento, certificando-se em de sobrepor as digitalizações por uma quantidade grande o suficiente para garantir que haja *features* comuns suficientes. Um problema maior está no processo de alinhamento de pares, que a medida que você alinha mais e mais pares, a solução pode derivar para que não seja mais globalmente consistente [9].

Já na sumarização e a compactação de vídeo, uma aplicação interessante de costura de imagem é a capacidade de resumir e comprimir vídeos tirados com uma câmera panorâmica. Embora o *stitching* de vídeo seja uma simples generalização de costura de várias imagens, a presença potencial de grandes quantidades de movimento independente, zoom da câmera e o desejo de visualizar eventos dinâmicos impõe desafios adicionais [9]. Além disso, o vídeo também pode ser uma fonte interessante de conteúdo para a criação de panoramas tirados de câmera em movimento. Outros movimentos de câmera e de superfícies de composição usando o conceito de mosaicos em variedades adaptativas e usado para gerar estereograma⁵ panorâmicos [9].

Uma outra variante presente são os panoramas baseados em vídeo são os mosaicos concêntricos. Ao invés de produzir uma única imagem panorâmica, o vídeo original completo é mantido e usado para re-sintetizar vistas usando o remapeamento de raios [9].

4 Alinhamento global

Algoritmos para alinhar imagens e o *stitching* para emendar foto-mosaicos estão entre os mais antigos e os mais usados na visão computacional. Em meados da década de 1990, as técnicas de alinhamento de imagens começaram a ser usada na construção de panoramas de ângulo aberto sem *stitching*, a partir de câmeras de mão. Assim devemos primeiro determinar o modelo matemático apropriado relacionado às coordenadas de pixel em uma imagem para coordenadas de pixel em outra [9].

Na maioria das aplicações recebemos mais do que um único par de imagens para registrar e seu objetivo é encontrar um conjunto globalmente consiste de parâmetros de alinhamento que minimizem o erro de registro entre todos os pares de imagens [9]. À seguir depois de realizar o alinhamento global, precisamos realizar os ajustes locais, como a remoção de paralaxe, para reduzir imagens duplas e borrar devido a erros de registro local, e por fim, iremos reconhecer panoramas, em que nos

⁵Representação planar que imita as diferenças entre as duas imagens retiniais de um objeto em profundidade [1].

for dado um conjunto de imagens não ordenadas para registrar, precisamos descobrir quais imagens vão juntas para formar um ou mais panoramas.

4.1 Ajuste de pacote

Uma maneira de registrar um grande número de imagens é adicionar novas imagens ao panorama, alinhando a imagem mais recente com as anteriores já existentes na coleção e descobrindo, se necessário, quais imagens sobrepor. Quando os panoramas são 360°, o erro acumulado pode levar à presença de lacuna (ou sobreposição excessiva) entre as duas extremidades do panorama, e pode ser corrigido com o fechamento de lacunas que estica o alinhamento de todas as imagens. Uma maneira é alinhar simultaneamente todas as imagens usando uma estrutura de mínimos quadrados para distribuir corretamente quaisquer erros de registro incorreto. Esse processo de ajuste de pacotes é chamado de ajuste simultâneo de parâmetros de pose para uma grande coleção de imagens sobrepostas [9].

4.2 Remoção do paralaxe

Após otimizado as orientações globais e as distâncias focais de nossas câmeras, poderemos descobrir que as imagens ainda não estão perfeitamente alinhadas, ou seja, a imagem *stitching* resultante parece borrada ou existem "fantasmas" em alguns lugares. Isso é causado por uma distorção radial não modelada, paralaxe 3D, que são pequenos movimentos de cena como acenar galhos de árvores e movimentos de cena em larga escala, pessoas entrando e saindo fora de fotos, porém cada um desses problemas pode ser tratado com uma abordagem diferente [9]. A seguir, é citados algumas soluções para algumas causas que foram ditas anteriormente.

A paralaxe 3D pode ser tratada fazendo um ajuste completo do pacote 3D, substituindo a equação de projeção, que modela as translações da câmera. Com relação aos movimentos, quando o movimento na cena é muito grande, ou seja, quando os objetos aparecem e desaparecem completamente, uma solução é selecionar *pixels* de apenas uma imagem de cada vez como a fonte para o composto final. Já com movimentos pequenos, o fluxo óptico pode ser usada para realizar uma correção apropriada antes da mesclagem usando o alinhamento local [9].

4.3 Reconhecendo panoramas

O reconhecimento de panoramas é o último passo para reconhecer quais imagens se encaixam para realizar o *stitching* da imagem totalmente automatizada. Caso o usuário tire as imagens em sequência, para que cada imagem sobreponha sua predecessora e também especifique as primeiras e últimas imagens a serem *stitching*, o ajuste de pacotes combinado com o processo de inferência de topologia pode ser usado para montar automaticamente em um panorama, porém os usuários costumam se movimentar em que pode iniciar uma nova linha em cima de uma anterior, necessário descobrir sobreposições de ponta a ponta [9].

Para realizar o reconhecimento é necessário encontrar primeiro todas as sobreposições de imagens emparelhadas usando um método baseado em características, após isso encontrar componentes conectados no grafo de sobreposição para "reconhecer" panoramas individuais. Porém, o mais difícil de se trabalhar com algoritmo de costura totalmente automatizado é decidir quais pares de imagens realmente correspondem às mesmas partes da cena. Estruturas repetidas podem levar a correspondências falsas ao usar uma abordagem baseada em *features*. Para amenizar isso, pode ser realizada uma comparação direta baseada em *pixels* entre as imagens registradas para determinar se elas são visões diferentes da mesma cena [9].

4.4 Direto vs. alinhamento baseado *feature*

Os primeiros métodos baseados em *features* ficavam confusos em regiões que eram muito texturizadas ou não tinham textura suficiente. Essas *features* eram distribuídas de forma desigual sobre as imagens, não correspondendo aos pares de imagens que deveriam estar alinhados [9].

Atualmente, os esquemas de detecção e correspondências de recursos são robustos e podem até ser usados para reconhecimento de objetos conhecidos a partir de visualizações amplamente separadas. Além disso, se as *features* forem distribuídas sobre a imagem e os descritores projetados repetidamente, geralmente podem ser encontradas correspondências suficientes para permitir a costura de imagem [9].

Já com relação a abordagem direta, sua desvantagem seria quando o alinhamento é baseado em *pixels* elas têm um alcance limitado de convergência. Essa abordagem pode ser utilizada em estrutura hierárquica, porém na prática é difícil usar mais de dois ou três níveis de uma pirâmide antes que detalhes importantes comecem a ficar desfocados, um caso onde esse tipo de abordagem funciona é em correspondências de quadros sequências em um vídeo. Já para corresponder parcialmente a sobreposição de imagens em panoramas baseados em fotos ou coleções de imagens em que o contraste ou o conteúdo variam muito, eles falham com muita frequência, por isso a primeira abordagem é mais preferida [9].

5 Composição

Após registrar todas as imagens de entrada em relação umas às outras precisamos decidir como produzir a imagem final do mosaico *stitching*. Essa etapa envolve a seleção de uma superfície de composição final e vista, a seleção de quais *pixels* contribuem para a composição final e como misturar de forma ideal esses *pixels* para minimizar costuras visíveis, borrões e "fantasmas". Além disso, essa etapa pode ser aplicado em foto montagem, já que o *stitching* da imagem é usada para compor fotografias parcialmente sobrepostas, mas também pode ser usada para compor tomadas repetidas de uma cena tirada com o objetivo de obter uma melhor composição e aparência de cada elemento [9].

5.1 Escolhendo a composição da superfície

A primeira escolha é como representar a imagem final. Caso apenas algumas imagens forem *stitching*, uma abordagem natural é selecionar uma das imagens como referência, e depois distorcer todas as outras imagens em seu sistema de coordenadas de referência. A composição resultante é às vezes chamado de panorama plano, uma vez que a projeção na superfície final ainda é uma projeção em perspectiva, já que as linhas retas permanecem [9].

Após escolher a parametrização de saída, é necessário determinar qual parte da cena será centralizada na visualização final, para uma composição final podemos escolher uma das imagens como referência, uma escolha razoável seria aquela que é geometricamente central. Quando os panoramas são maiores (cilíndricos ou esféricos), pode se escolher um subconjunto da esfera de visualização que tenha sido visualizado. Já em panoramas 360°, poderia ser a imagem do meio a partir da sequência de entradas, ou a primeira imagem, assumindo que ela contém o objeto de maior interesse [9].

Depois disso, é necessário calcular o mapeamento entre as coordenadas dos *pixels* de entrada e saída. Se a superfície da composição final é plana e as imagens de entrada não tem distorção radial, a transformação de coordenadas é a homografia simples. Caso a superfície composta final tiver outra forma analítica, precisamos converter cada pixel no panorama final em um raio de visualização e mapear de volta em cada imagem de acordo com a projeção. Enfim, caso a superfície da composição final ser um poliedro mapeado por textura, um algoritmo mais sofisticado deve ser usado. Não apenas as coordenadas do 3D, mapa de texturas devem ser manuseadas adequadamente, mas uma pequena quantidade de *overdraw* fora do *footprint* do triângulo é necessária, para garantir que os *pixels* de textura que estão sendo interpolados durante a renderização tenham valores válidos [9].

É possível que o passo anterior, gere os endereços de *pixels* corretos, porém é necessário prestar atenção aos problemas de amostragem. O problema que pode ocorrer é calcular o pré-filtro apropriado, que depende da distância entre amostras vizinhas em uma imagem de origem [9].

5.2 Seleção de *pixels* e ponderação (desconfiguração)

Depois de ter mapeados na superfície, é necessário decidir como mesclar criando um panorama atraente. Se as imagens estiverem em perfeito registro e expostas de forma idêntica, qualquer combinação de *pixel* ou combinação serve. Porém, imagens reais, *stitching* visíveis, podem ocorrer desfoque ou "fantasmas". Para criar panoramas limpos e agradáveis envolve decidir quais *pixels* usar e como aumentar ou combinar [9].

A maneira mais simples de criar uma composição final é pegar o valor médio de cada pixel, porém ela não funciona muito bem, já que as diferenças de exposição, os registros incorretos e o movimento da cena são muito visíveis. Caso os objetos se movam rapidamente seja o único problema, pode-se

utilizar um filtro mediano para remover. Mas a melhor abordagem para calcular a média é pesar mais os *pixels* próximos ao centro da imagem e diminuir os *pixels* próximos às bordas. Outro tipo de média seria a média ponderada com um mapa de distância que é chamada de difusão e faz um trabalho razoável de mistura com diferenças de exposição, porém o embaçamento e "fantasmas" ainda podem ser problemas [9].

A computação do diagrama de Voronoi é uma maneira de selecionar as junções entre as regiões onde diferentes imagens contribuem para a composição final, porém elas ignoram totalmente a estrutura da imagem local subjacente à costura. Uma abordagem melhor é colocar as emendas nas regiões onde as imagens estão de acordo, de modo que as transições de uma fonte para outra não sejam visíveis, assim evita cortar objetos em movimentos [9].

5.3 Misturando

Depois que as junções entre as imagens tiverem sido determinadas e os objetos indesejados removidos, é necessário mesclar as imagens para compensar as diferenças de exposição e outros desalinhamentos. Porém, é difícil de obter um equilíbrio agradável entre suavizar as variações de exposição de baixa frequência e manter transições nítidas o suficiente para evitar o desfoque [9].

Uma solução para resolver esse problema é a pirâmide laplaciana. Em vez de usar uma única largura de transição, é utilizada uma largura adaptável à frequência, criando uma pirâmide de passagem de faixa e fazendo as larguras dentro de cada nível uma função do nível. Uma outra solução é a mesclagem de imagens de várias bandas que executa as operações no domínio de gradiente [9].

A combinação de domínio de pirâmide e gradiente pode compensar a quantidade moderada de diferenças de exposição entre as imagens, porém essa diferença é grande, abordagens alternativas são necessárias. Enfim, a maneira mais correta de lidar com as diferenças de exposição é costurar imagens no domínio de radiância, ou seja, converter cada imagem em radiância usando seu valor de exposição e criar uma imagem de alta faixa dinâmica [9].

References

- [1] Herman Martins Gomes. Intro estereo, 2015.
- [2] Robert Karol Harrison Chau. Robust panoramic image stitching, January 2015.
- [3] Derek Hoiem. Image stitching, October 2010.
- [4] Daeho Lee Jaeyoung Yoon. Real-time video stitching using camera path estimation and homography refinement, December 2017.
- [5] Oubong Gwun Luo Juan. Surf applied in panorama image stitching, July 2010.
- [6] Bryan Morse. Image stitching, 2012.
- [7] Gerhard Roth. Homography, February 2014.
- [8] Richard Szeliski. Image alignment and stitching, January 2005.
- [9] Richard Szeliski. *Computer Vision: Algorithms and Applications*. Springer, September 2010.
- [10] Raquel Urtasun. Computer vision: Panorama, February 2013.
- [11] Jianxiong Xiao. Problem set 3: Stitching a panorama, 2014.