

Laboratório VISGRAF

Instituto de Matemática Pura e Aplicada

Zoom para Realidade Virtual
Baseado em Transformação de Moebius

Andre Potengy
Luiz Velho

Technical Report TR-17-09 Relatório Técnico

November - 2017 - Novembro

The contents of this report are the sole responsibility of the authors.
O conteúdo do presente relatório é de única responsabilidade dos autores.

Zoom para Realidade Virtual Baseado em Transformação de Moebius

André Potengy

VISGRAF Lab, Instituto de Matemática Pura e Aplicada

potengy@impa.br

Luiz Velho

VISGRAF Lab, Instituto de Matemática Pura e Aplicada

lvelho@impa.br

Abstract

O interesse em imagens e vídeos omnidirecionais vem crescendo com o aumento da disponibilidade de dispositivos de visualização e câmeras 360° de alta definição, desde o YouTube em smartphones até os sofisticados Head Mounted Displays que estão cada vez mais ao alcance do público. Como características em destaque em vídeos exibidos em realidade virtual, a liberdade do espectador em escolher para onde olhar e a excepcional sensação de imersão, por um lado oferecem enormes possibilidades de autoria para os criadores de conteúdo e por outro impõem restrições ao uso de elementos tradicionais da linguagem cinematográfica. Entre estes recursos, existe um debate sobre a aplicabilidade do zoom em VR, seja pela falta de qualidade do resultado das transformações convencionais em imagens esféricas, seja pelo risco de quebra da imersão.

Neste trabalho abordamos o zoom como elemento de linguagem cinematográfica aplicável em realidade virtual. Descrevemos a transformação de Moebius em vídeos como a formulação matemática mais apropriada para operar em funções na esfera e que consequentemente resulta na ferramenta de zoom natural para edição e visualização de imagens omnidirecionais. Além disso mostramos diferentes maneiras de utilizar este zoom em tempo de exibição, de acordo com o ponto de vista do espectador e parâmetros definidos pelo criador de conteúdo em tempo de produção.

Introdução

Com o avanço e popularização das tecnologias de captura e visualização, aumenta a relevância de compreendermos como os recursos utilizados da linguagem cinematográfica se aplicam nesta nova realidade, que recursos não fazem mais sentido e que novos recursos podem ser criados.

Em VR, os principais pilares que caracterizam uma experiência são as oportunidades de imersão, o potencial para uma experiência única de empatia e o controle pela audiência [7]. São estes os três principais elementos que observamos como oportunidades e também como condicionantes ao uso de recursos do cinema e do teatro na produção de trabalhos em VR.

Um dos recursos em debate pela sua aplicação em Cinematic VR é o Zoom. De um lado, há o argumento que o zoom não se aplica ou porque os movimentos de câmera devem ser de domínio do espectador ou porque o zoom na esfera distorce a geometria do que está sendo observado, em ambos os casos, interrompendo a imersão. No nosso trabalho estamos desenvolvendo experimentos que abordam estes dois aspectos: a qualidade do zoom e a forma ou dinâmica de aplicação.

Em relação a qualidade, desenvolvemos um zoom baseado nas transformações de Moebius na esfera. Por se tratarem de transformações conformes, preservam características da geometria da imagem em transformação, de forma que o resultado parece mais natural. Esta percepção de naturalidade depende da faixa de ângulo de visão, uma vez que a transformação de Moebius distorce mais a imagem quanto maior for a distância do ponto transformado até o centro de referência ou ponto de vista. Adicionalmente, o movimento de cabeça durante a aplicação de zoom gera uma percepção de não correspondência com o movimento da imagem, semelhante ao movimento durante o uso de um binóculo. Outro fenômeno a se observar é que a aplicação do zoom, se exagerada, corta a naturalidade da experiência e a imersão.

Por estes motivos, entendemos que devem ser estudados os limites em que o zoom seja suficiente para proporcionar uma experiência natural e ao mesmo tempo proporcionar o efeito dramático desejado pelo criador de conteúdo.

Trabalhos Relacionados

O trabalho apresentado em [1] descreve um sistema completo em tempo real para aplicar zoom e pan interativos em vídeos panorâmicos de alta definição. Comparado aos sistemas então existentes usando panoramas em perspectiva com recorte, a abordagem em [1] cria um panorama cilíndrico. Neste caso, a perspectiva é corrigida em tempo real, e o resultado é um zoom melhor e mais natural.

Em [2], Peñaranda, Luiz Velho e Sacht apresentam uma técnica para melhorar a percepção de qualidade de visualização panorâmica. Os principais ingredientes desta abordagem são, de um lado, considerar a esfera de visão como sendo a esfera de Riemann, o que torna natural a aplicação de transformações de Moebius sobre a imagem de entrada, e, por outro lado, o esquema de projeção que muda de função conforme o campo de visão usado. Também introduzem uma implementação do método, comparam com imagens produzidas com outros métodos e mostram que as transformações podem ser feitas em tempo real, o que torna esta técnica atraente para aplicações panorâmicas interativas existentes ou novas.

Schleimer e Segerman propõem as transformações de Moebius como ferramentas naturais de escala e rotação a serem usadas na edição de imagens esféricas em [3]. Como aplicações, eles mostram como obter o "Droste" e outros efeitos visuais interessantes usando Moebius e outras transformações conformes.

Em [4], Souto, Sacht e Luiz Velho apresentam um novo método para transformar imagens omnidirecionais com base em uma combinação de transformações de Moebius no plano complexo que são localizadas por pesos Gaussianos para limitar a ação destes mapeamentos a regiões de interesse. Como nenhuma otimização ou método numérico estão envolvidos, é possível atingir desempenho em tempo real com o método, tornando-o aplicável a serviços de streaming. Os resultados mostram o potencial desta técnica em áreas como segurança e arte em geral.

A Linguagem Cinematográfica e a Realidade Virtual

Chamamos de linguagem cinematográfica o conjunto de planos, ângulos, movimentos de câmera, recursos de montagem, luzes, sons, argumento, enquadramentos e sequências que compõem o universo de um filme. Para isso, é preciso ter em conta que cada um tem seu efeito psicológico, um valor dramático específico e exerce seu papel dentro da totalidade que é um filme.

Comparando os suportes de mídia para o cinema tradicional e o cinema 360 graus, temos no primeiro caso a televisão e filme e no segundo caso projeções esféricas em domos e dispositivos com rastreamento da direção de visada do observador, tais como os Head Mounted Displays (HMD) e tablets. Ver Figura 1.

- Conventional Cinema
 - HD Television
 - Theater / Panavision
- 360 Degrees Cinema
 - Omnimax / Full Dome
 - HMD
 - Tablet

Fig. 1 - Cinema Tradicional versus cinema 360 graus.

Ao analisarmos os elementos da linguagem tradicional do cinema em comparação com suas possibilidades em realidade virtual, além dos efeitos dramáticos e psicológicos intencionais, devemos considerar os efeitos indesejáveis que podem causar sobre o espectador durante a sua experiência.

Se empregados de forma inapropriada, alguns recursos podem quebrar a sensação de imersão, por enviar uma “mensagem” ao espectador de que ele não está fisicamente presente naquela experiência. Este “despertar para a realidade”, além de anular o principal efeito do VR, a imersão, pode causar problemas como desconforto visual (*motion sickness*).

Outra característica fundamental em VR a ser considerada é a liberdade do espectador em olhar para onde desejar. Enquanto essa liberdade proporciona enormes oportunidades na experiência VR, ao mesmo tempo interfere o poder do diretor na condução do olhar do espectador. Na cinematografia tradicional o diretor determina o que o espectador vai assistir, sob qual plano, ângulo, enquadramento e movimentação de câmera, enquanto em VR o diretor se restringe a escolher a posição em que o espectador está em relação a cena, buscando sutileza nos efeitos visuais para não comprometer a imersão.

Ainda sob o risco de quebra de imersão, os recursos de montagem são severamente restritos, uma vez que cada corte envia um sinal ao espectador de que aquela experiência que está vivendo não é real. Com isso, os autores ou buscam disfarçar os recursos de montagem na narrativa ou buscam minimizar a sua utilização.

Os movimentos de camera constituem a base da linguagem cinematográfica tradicional, por isso se torna importante analisa-los na comparação com a linguagem cinematográfica para realidade virtual.

Cinema Tradicional

- ❖ **Movimentos da Câmera Cinematográfica:** Constituem a base técnica do plano em movimento. São definidos levando-se em conta se o movimento da câmera é de translação (locomovendo-se em avanço ou recuo, subindo ou descendo), de rotação (em torno do seu eixo) ou de mudança do campo de visão relativo ao espectador. Ver Figura 2.

- **Travelling:** A câmera é movida sobre um carrinho (ou qualquer suporte móvel, como a grua) num eixo horizontal / vertical e paralelo ao movimento do objeto filmado. Este acompanhamento pode ser lateral ou frontal, neste último caso podendo ser de aproximação ou de afastamento.
- **Panorâmica:** A câmera se move em torno do seu eixo, fazendo um movimento giratório, sem sair do lugar. Trata-se de um movimento da câmera que pode ser horizontal (da esquerda para a direita ou da direita para a esquerda), vertical (de cima para baixo ou vice-versa) ou oblíquo. A panorâmica vertical é também conhecida como tilt.
- **Zoom:** No zoom, a câmera se mantém fixa e é seu conjunto de lentes que se move, fazendo com que o objeto se apresente mais afastado ou mais próximo na imagem.

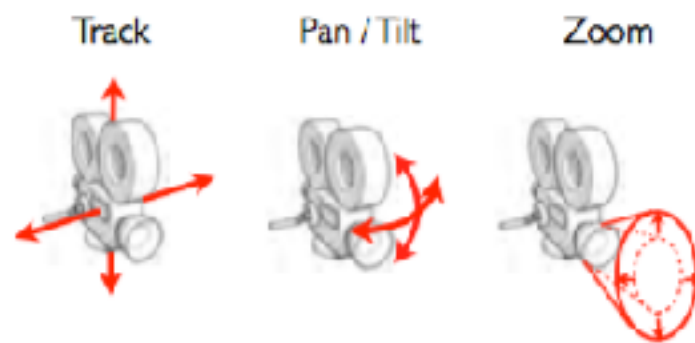


Fig. 2 - Movimentos da Camera Tradicional

Cinema 360 graus

- ❖ **Movimentos da Câmera Omnidirecional:** Como em VR o espectador tem o controle do ponto de vista, ao contrário do cinema tradicional, no cinema 360 graus os movimentos da camera podem ser determinados no momento da captação da imagem pelo diretor ou durante a visualização pelo observador. Ver Figura 3.
- **Travelling:** O movimento de translação pode ser aplicado em tempo de gravação, em geral nas situações dramáticas que supõe o espectador se deslocando, por exemplo em um veículo.
- **Panorâmica:** Já o movimento de rotação é realizado em tempo da visualização por iniciativa do observador ou por transformação da imagem.
- **Zoom:** Existe debate sobre a utilização ao não do zoom em VR. Em primeiro lugar está a premissa de que as câmeras omnidirecionais não tem recursos para alterar o seu campo de visão, um vez que registram uma imagem de 360 graus. Entretanto, de fato o zoom resulta da variação do campo de visão na cena visualizada e o campo de visão de referência do espectador. Assim é possível criar a sensação de zoom em VR com uma transformação da imagem. Por outro lado, existe a dificuldade em se desenvolver uma transformação de zoom que seja percebida como natural para o observador, dado que a superfície de projeção é uma esfera e não um plano. Além disso, mesmo dispondo de um zoom natural, a câmera não se mantém fixa. Como o observador tem a liberdade de olhar para qualquer direção, ele pode não estar olhando para o objeto de interesse para o zoom.

Neste trabalho, oferecemos o zoom utilizando transformação de Moebius como uma solução que cria uma percepção natural para o observador e também exploramos formas de aplicar o zoom com controles de tempo de edição e de tempo de exposição.

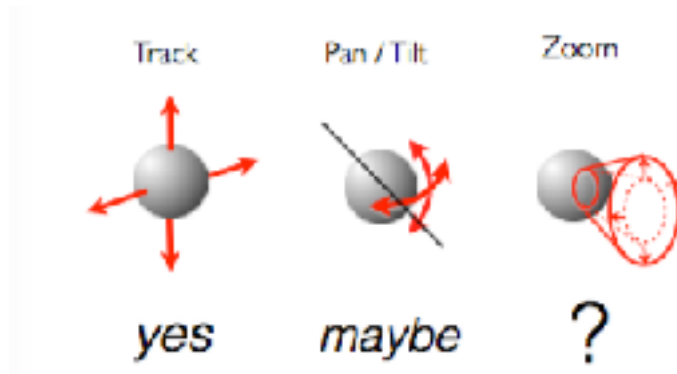


Fig. 3 - Movimentos da Camera Omnidirecional

Transformações de Moebius para Realidade Virtual

De forma quase universal, as imagens e vídeos esféricos são armazenados e transmitidos por meio de projeção equiretangular: pontos na esfera são dados pela latitude e longitude [2] e a imagem completa é armazenada como uma imagem retangular. Este formato é conveniente para a atual infraestrutura de processamento de imagens. Porém, há um problema ao se editar imagens esféricas, o software desenvolvido para imagens retangulares ordinárias não conhece as projeções equiretangulares. Não se aplicam nem a rotação nem a escala.

Por ser conforme, a transformação de Moebius apresenta a vantagem de preservar os ângulos e produzir outros efeitos interessantes, favorecendo uma experiência mais natural do observador na aplicação de escala ou zoom.

As transformações de Moebius podem ser classificadas de forma canônica em três tipos, a saber: Elíptica; Hiperbólica e Parabólica. Ver Figura 4.

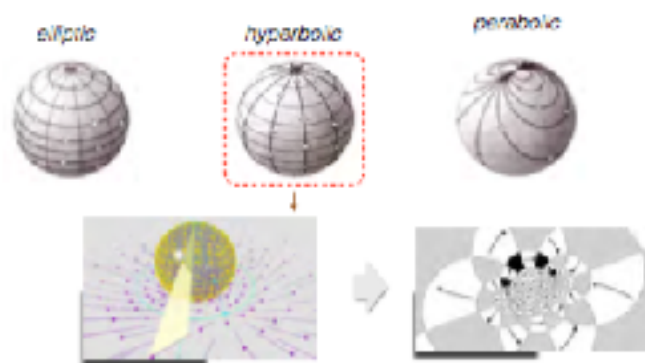


Fig 4. Transformações Canônicas de Moebius

As transformações de Moebius são definidas como mapeamentos do plano complexo estendido e que corresponde à esfera de Riemann. Por esse motivo, essas transformações constituem a base matemática natural para operar com imagens omnidirecionais. Além disso, podemos constatar que as transformações canônicas de Moebius tem uma correspondência direta com os movimentos de câmera em VR, exceto pela translação da câmera a qual não pode ser efetuada por meio de transformação de imagem. Ou seja, o movimento panorâmico

corresponde à transformação elíptica enquanto o zoom corresponde à transformação hiperbólica. Pode-se conjecturar que a transformação parabólica afeta relações de perspectiva. Ver Figura 5.

- Omnidirectional Images and Moebius Transformations
- Pan / Tilt ↔ Elliptic Transform
- Zoom ↔ Hyperbolic Transform
- Perspective ↔ Parabolic Transform ?

Fig. 5 - Correspondência entre movimentos de camera e transformações de Moebius

Representações de Imagens Omnidirecionais

Toda a cena observada a partir de um ponto de vista fixo é modelada como uma esfera unitária centralizada no ponto de vista ($\{S^2 = (x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$), onde cada ponto na esfera possui uma cor associada, a cor que é vista quando alguém olha através deste ponto. Por conveniência, assumimos que a origem no \mathbb{R}^3 é o ponto de vista.

Esta esfera colorida é chamada *esfera visível* ou *imagem omnidirecional*. Nota-se que a esfera visível representa completamente o campo de visão com 360° de longitude por 180° de latitude.

Uma representação bem conhecida e útil da esfera S^2 é pelas coordenadas de latitude e longitude, ou formato equiretangular:

$$r: [-\pi, \pi] \times \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \rightarrow S^2$$

$$(\lambda, \phi) \leftrightarrow (\cos\lambda \cos\phi, \sin\lambda \cos\phi, \sin\phi)$$

Esta representação está ilustrada na Figura 6:

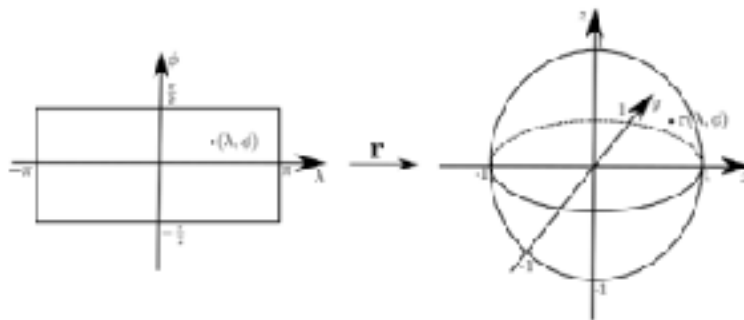


Fig. 6: Representação r em latitude/longitude

É conhecido que um plano complexo pode ser identificado como uma esfera unitária (no nosso caso, a esfera visível) por meio de uma projeção estereográfica. Identificando o ponto a partir do qual as linhas de projeção emanam como ∞ , obtemos um mapeamento contínuo $SP: S^2 \rightarrow C_\infty = C \cup \infty$ que tem uma inversa contínua $SP^{-1}: C_\infty \rightarrow S^2$. Uma propriedade importante da projeção estereográfica é sua conformidade, ou seja, ela preserva ângulos e

mapeia círculos em círculos ou linhas. Esta forma de interpretar o plano complexo estendido C_∞ como uma esfera foi formulado por Riemann e recebe o nome de *Esfera de Riemann*. No nosso trabalho, consideramos a esfera visível como sendo a esfera de Riemann com o ∞ sendo o ponto oposto ao plano de projeção.

Transformação de Moebius

Utilizamos modelo semelhante ao utilizado em [4].

Definição 1. O mapeamento $S: C_\infty \rightarrow C_\infty$ na forma $S(z) = \frac{az + b}{cz + d}$ é chamada uma transformação linear fracionária. Se a, b, c e d satisfazem a $ad - bc \neq 0$ então S é chamada uma *Transformação de Moebius*.

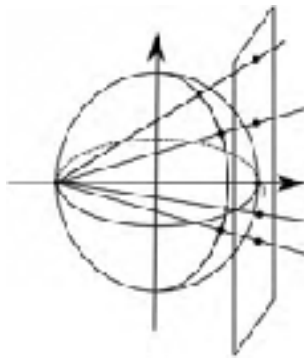


Figura 7: Projeção Estereográfica

Computação da Transformação de Moebius

A computação da transformação de Moebius para imagens omnidirecionais representadas no formato equirectangular pode ser efetuada por uma sequência de passos com descrito a seguir.

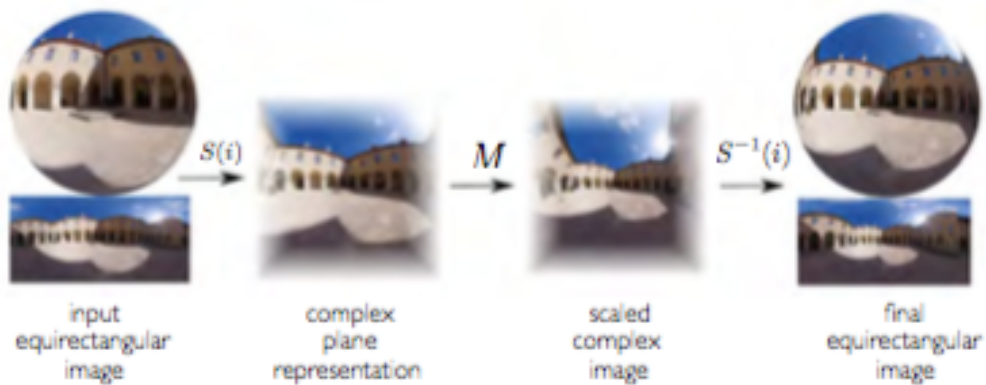
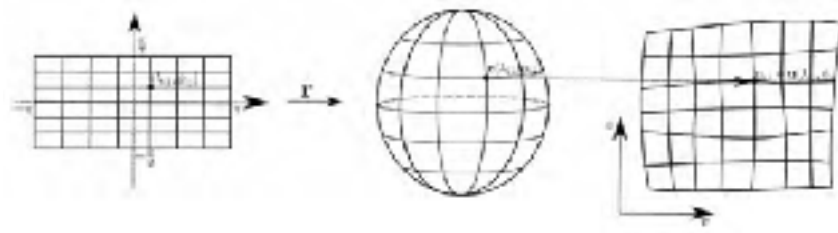


Fig. 8 - Sequência de Implementação da Transformação de Moebius.

Nessa sequência a imagem é mapeada da parametrização latitude / longitude para a esfera de Riemann e depois para o plano complexo estendido onde a transformação se realiza. Subsequentemente, a inversa do mapeamento é aplicada.

Esses passos estão especificados a seguir.

Warping:



$$\mathbf{u}: S^2 \rightarrow \mathbb{R}^2$$

$$(\lambda, \phi) \mapsto (u(\lambda, \phi), v(\lambda, \phi))$$

Alinhamento do eixo Z com o eixo da câmera:

Seja \mathbf{M}^T a matriz transposta da matriz de transformação correspondente ao translado, rotação e escala da câmera.

$$\mathbf{A}: S^2 \rightarrow S^2 \setminus \{(0,0, -1)\}$$

$$(x_0, y_0, z_0) \mapsto (x, y, z) = \mathbf{M}^T \times (x_0 \ y_0 \ z_0 1)$$

Projeção estereográfica sobre o plano complexo:

$$\mathbf{S}: S^2 \setminus \{(0,0, -1)\} \rightarrow \mathbb{C}$$

$$(x, y, z) \mapsto (u, v) = \left(\frac{2x}{z+1}, \frac{2y}{z+1} \right)$$

Forma Polar:

$$u + iv = (u, v) \mapsto (r, \theta) = (\sqrt{u^2 + v^2}, \arctan2(u, v)) = re^{i\theta}$$

No caso do zoom, a transformação Hiperbólica corresponde à Multiplicação Complexa:

$$(r, \theta) \mapsto (\tilde{r}, \tilde{\theta}) = (rs, \theta + \alpha)$$

Sendo $s \in \mathbf{R}$ a escala, i.e., o zoom e $\alpha \in \mathbf{R}$ o ângulo de rotação da transformação de Moebius.

Forma Cartesiana:

$$(\tilde{r}, \tilde{\theta}) \mapsto (\tilde{u}, \tilde{v}) = (\tilde{r} \cos \tilde{\theta}, \tilde{r} \sin \tilde{\theta})$$

Inversa da estereográfica (do plano complexo para a esfera):

$$\mathbf{S}^{-1}: C \rightarrow S^2 \setminus \{(0,0, -1)\}$$
$$(\tilde{u}, \tilde{v}) \mapsto (x_f, y_f, z_f) = \left(\frac{\tilde{u} \times (1 + z_f)}{2}, \frac{\tilde{v} \times (1 + z_f)}{2}, \frac{\tilde{u}^2 + \tilde{v}^2 - 4}{\tilde{u}^2 + \tilde{v}^2 + 4} \right)$$

Desfaz o alinhamento do eixo Z com a câmera, retornando a esfera ao estado inicial de rotação:

$$\mathbf{A}^{-1}: S^2 \setminus \{(0,0, -1)\} \rightarrow S^2$$
$$(x_f, y_f, z_f) \mapsto (x_m, y_m, z_m) = \mathbf{M} \times (x_f \ y_f \ z_f 1)$$

Implementação em Tempo Real na GPU

Para processar a transformação de Moebius em tempo real em videos ou mesmo em imagens panorâmicas, desenvolvemos um shader que tira proveito do hardware gráfico do equipamento utilizado para a visualização (GPU). Este shader é associado ao render da esfera.

Utilizamos o game engine Unity e o HMD Gear VR como plataformas para os nossos experimentos. Tipicamente, implementamos a visualização de videos ou imagens panorâmicas em VR criando uma cena no Unity, com uma esfera e uma camera centradas na origem da cena. O ponto de vista do observador é definido em tempo de execução pela rotação do HMD ao qual a camera é associada.

No caso de videos, também incluímos na cena um objeto do Unity chamado Video Player que controla a projeção do video pelo render da esfera.

O shader processa os vértices da malha da esfera, retornando uma nova posição para cada vértice, resultante da transformação de Moebius. Sobre esta malha resultante, o Unity mapeia as imagens, aplicando as distorções correspondentes.

A transformação de zoom pode ser definida em relação à uma direção na imagem panorâmica ou com referência a linha de visada do observador.

No primeiro caso, a transformação é calculada em um ponto fixo da esfera, o nosso shader recebe como parâmetros os valores de escala e rotação a serem aplicados na transformação de Moebius e a imagem panorâmica.

Para cada vértice da malha do objeto, o shader aplica os cálculos descritos anteriormente e retorna os vértices transformados. As cores que o shader retorna são aquelas encontradas na posição correspondente na imagem panorâmica.

Neste caso, o observador percebe o efeito natural de zoom e rotação quando seu eixo de visão está alinhado com o eixo usado como referência para a transformação. Como a transformação de escala é aplicada em toda a superfície da esfera, no ponto de referência da transformação se percebe um zoom natural, porém quanto mais distante deste ponto, maior é a distorção percebida pelo observador.

No segundo caso, para que o ponto de referência da transformação coincida com o ponto de vista do observador, optamos por calcular uma matriz de rotação e passar esta matriz para o shader como parâmetro de entrada. O shader usa esta matriz para aplicar uma rotação nos vértices, alinhando o eixo de visão do observador com o eixo de referência dos cálculos da transformação de Moebius, para então aplicar estes cálculos. Fazendo isso, o efeito de zoom é percebido sempre no ponto de vista do observador.

A imagem utilizada foi capturada por uma câmera Lady Bug da Point Grey. Para facilitar a depuração dos algoritmos de zoom em tempo de exibição, incluímos marcadores na imagem utilizando o Adobe Photoshop e também um HUD virtual indicando os valores de zoom e das coordenadas do ponto de vista e uma cruz indicando a posição do ponto de vista na visualização.

Testes de Percepção com o Zoom

Conforme mencionamos, entendemos que deve haver limites para os valores de escala no zoom, dependendo da abertura do ângulo de visão que cada HMD oferece. Nos nossos testes vamos determinar estes limites por meio da utilização do zoom com uma variedade de pessoas e HMDs.

Como o zoom em VR é uma transformação global na esfera de visualização, a ampliação ou redução da região observada também distorcem o restante da imagem projetada da esfera. Sendo a distorção tanto maior quanto mais se afasta do ponto de vista, no centro da transformação. Em um zoom-out, ao se aumentar o efeito e afastamento, chega um momento em que se enxerga distorções da geometria dos objetos em cena. Já no zoom-in, felizmente a região distorcida começa fora e fica progressivamente mais distante do campo de visão ao se aumentar o efeito de aproximação.

Uma vez estabelecidos os limites que previnem a percepção de distorção das imagens, também devemos determinar a faixa de valores em que o zoom cause um efeito narrativo e que ao mesmo tempo não seja conscientemente percebido pelo observador. Este recurso pode ser utilizado, por exemplo, para realçar a reação de uma personagem, ou mesmo para direcionar a atenção do espectador para um determinado objeto ou região da cena, no momento em que ele olha para este objeto ou região de interesse. O destaque deste objeto ou região gerado pelo zoom pode atrair a atenção e o foco do espectador, mas se for exagerado, o espectador pode perceber o movimento artificial da imagem, interrompendo a sua imersão.

Outro fenômeno que devemos tratar é a sensação de movimento quando o zoom está fora do valor neutro. Tanto para o zoom-in como para o zoom-out, ao movimentar a cabeça, o espectador experimenta uma sensação de que a rotação do mundo a sua volta está mais rápida (no zoom-in) ou mais lenta (no zoom-out) do que o normal. É a mesma sensação que se tem ao usar um binóculo na posição normal ou invertida e movimentar a cabeça. Também neste caso, devemos encontrar os limites de valores em que este fenômeno não seja perceptível.

Resultados dos Testes

Em nossos testes buscamos encontrar os valores de zoom em que as distorções de imagem permanecem fora do campo de visão, em que se percebe que o movimento da imagem corresponde ao movimento da cabeça e em que o espectador percebe conscientemente o efeito de zoom ao observar objetos destacados.

Percepção de distorções da imagem

Conforme descrito anteriormente, no zoom-in baseado na transformação de Moebius a imagem permanece íntegra independentemente dos valores do efeito de aproximação. O zoom-out a partir de uma visualização já processada pelo zoom-in também não apresenta distorções até se chegar no valor neutro (estado original da visualização).

Por outro lado, o zoom-out gera distorção periférica progressiva ao se aumentar o efeito de distanciamento a partir da visualização original. Estas distorções podem ser melhor notadas ao se observar o momento que linhas retas dentro do campo de visão começam a curvar. Em cenas mais “orgânicas”, em que faltem linhas retas, é possível que o zoom-out possa assumir valores maiores de distanciamento do que em cenas com presença de linhas retas.

Correspondência dos movimentos do espectador com a imagem

Em nossos testes percebemos que o efeito da movimentação do ponto de vista do observador quando aplicado o zoom pode não ser natural, dependendo do grau de escala aplicado. Sem o zoom, ou com o zoom em grau neutro, o movimento da cabeça do observador resulta em um movimento natural do ponto de vista, como se estivesse acompanhando o movimento da cabeça em sincronismo.

A partir de um determinado limite de valor de escala, seja para ampliar como para reduzir, o movimento da imagem já não acompanha naturalmente o movimento cabeça do observador, produzindo um efeito semelhante ao de se movimentar a cabeça utilizando um binóculo.

Para entendermos os limites naturais de escala, criamos uma aplicação com controle de zoom pelo observador e registramos os valores de zoom em que o usuário percebe o movimento da imagem como natural em relação ao movimento da cabeça.

Experimentos e Efeitos Narrativos

Realizamos também experimentos com finalidade de produzir efeitos para narrativa cinematográfica. Nesse sentido, um aspecto a ser considerado no zoom é a sua forma ou dinâmica de aplicação, que corresponde basicamente em como definir o valor da escala a ser usada no cálculo da transformação de Moebius.

Nos nossos experimentos criamos duas formas de definir o valor da escala a ser aplicada no zoom: controlada ou pré-definida. A escala controlada é aquela em que o observador define o valor da escala em tempo de execução, utilizando um controle vinculado ao HMD. A escala pré-definida é aquela em que o valor de escala para cada ponto da imagem é definido em tempo de produção.

O modelo em que a escala é controlada pelo observador nos parece pouco interessante do ponto de vista de narrativa. Este controle foi usado para nos ajudar a definirmos os limites de valores que discutimos anteriormente. Porém, entregar o controle do zoom ao observador representa um risco de quebra de imersão ou de distração e de não ter significado algum na narrativa, a não ser que, como parte da história, o observador vista um equipamento visual que tenha zoom controlado por ele.

No caso de escala pré-definida, implementamos dois experimentos: o primeiro, utilizando pontos de referência com valores de escala atribuídos previamente e o segundo, utilizando um mapa de zoom cujo tom de cinza em cada pixel corresponde aos valores de escala.

No primeiro caso, o valor da escala na posição em que estiver o ponto de vista é calculado pela média dos valores associados aos pontos de referência, ponderada pelos inversos das distâncias geodésicas na esfera do ponto de vista até estes pontos.

Já no caso do mapa de zoom, criamos uma imagem em que cada pixel possui um tom de cinza que representa o valor da escala naquela posição. Em tempo de exibição, calculamos qual seria a coordenada do pixel correspondente ao ponto de vista do observador na imagem e com isso definimos o valor da escala como uma função do tom de cinza daquela coordenada no mapa de zoom. Para criar o mapa de zoom, utilizamos o Photoshop e pintamos uma camada sobre a imagem panorâmica que de fato é visualizada. A resolução e as dimensões do mapa de zoom independem das dimensões e resolução da imagem a ser visualizada.

As duas abordagens apresentam vantagens e desvantagens dependendo da aplicação. A atribuição de valores a um conjunto de pontos é pouco intuitiva, mas pode facilitar uma eventual dinâmica de deslocamento destes pontos, no caso do objeto de interesse estar se movimentando. Já o mapa de zoom é confeccionado de forma muito mais natural, mas apresenta maior esforço no caso de o zoom ter de acompanhar um objeto em movimento, pois em vez de um mapa é necessário criar uma animação (sequência de mapas).

Uma aplicação especial para o zoom em VR é simular o efeito de travelling, possibilitando um aumento potencial na sensação de imersão por proporcionar a percepção de movimento da imagem correspondendo ao deslocamento da cabeça do espectador. Em VR, o centro da esfera de visualização acompanha a posição do HMD. Deslocar o centro da esfera de visualização em relação a posição do HMD cria problemas, principalmente em relação a distorção da imagem com o giro da cabeça. Então, como a esfera acompanha o movimento da cabeça do observador, o mesmo não altera a imagem projetada na esfera. Como consequência, uma pessoa se inclinando para ver um objeto mais de perto percebe a imagem, como se ele não tivesse chegado mais perto. Por não ser um fenômeno natural, também representa uma quebra de imersão.

Esta quebra de imersão pode ser solucionada aplicando um zoom cuja escala seja controlada pelo movimento de translação do HMD. Se o observador mover a cabeça para chegar mais “perto” de um objeto na cena, ao aumentarmos a escala no seu ponto de vista, criamos a ilusão de que o objeto ficou mais próximo. A restrição desta solução é justamente o limite em que a escala provoca uma sensação artificial de movimento de imagem ao girar da cabeça, como veremos adiante.

Como o Gear VR não captura movimento de translação, o efeito de deslocamento pode ser acionado por um controle manual, remetendo a uma experiência artificial. Já com o Oculus Rift, podemos ter uma captura moderada de movimento de translação de cabeça. No caso de um HTC Vive, em que é possível o usuário se movimentar em um espaço maior, o efeito de travelling se torna um desafio também maior a ser usado de forma restrita. Se a cena contém objetos de interesse relativamente distantes, a falta de paralaxe criaria um problema cognitivo.

Conclusão

Dentre os recursos da linguagem cinematográfica, exploramos o efeito de zoom em VR como uma possível ferramenta de apoio a narrativa. Sugerimos o zoom baseado na transformação e Moebius, abordando tanto os aspectos de qualidade da imagem resultante como as formas de dinâmicas de aplicação.

Consideramos como o zoom gerado a partir da transformação de Moebius é adequado para VR, produzindo resultados que são percebidos como naturais para o observador, respeitando-

se limites de valores de escala que dependem do tipo de aplicação que se deseja e também do tipo de HMD utilizado. Fora destes limites, as distorções geradas representam risco à qualidade da imagem e de outros fatores fundamentais a imersão, embora possam ser utilizados como efeitos visuais compatíveis com a narrativa.

Apresentamos duas formas de edição para definição de zoom em tempo de produção. Em ambos os casos, a transformação é aplicada globalmente em tempo real, acompanhando o ponto de vista do observador. A partir dos nossos testes, concluímos que é possível configurar mapas de zoom de forma a aplicar o efeito visual sem prejudicar a experiência do observador, obtendo o efeito dramático desejado, ainda que sutil.

Como oportunidades, conjecturamos que outras aplicações de efeitos visuais decorrentes da transformação de Moebius hiperbólica podem ser exploradas com eficácia na narrativa em VR. A aplicação de transformações diferentes em camadas de vídeo panorâmico é uma delas.

Referências

- [1] Vamsidhar Reddy Gaddam et al. *“Interactive Zoom and Planning from Live Panoramic Video”*. Proceedings of Network and Operating System Support on Digital Audio and Video Workshop. Singapore, Mar - 2014
- [2] Luis Peñaranda, Luiz Velho, Leonardo Sacht. *“Real-time Correction of Panoramic Images using Hyperbolic Möbius Transformations”*. Rio de Janeiro. IMPA. 2015
- [3] Saul Schleimer, Henry Segerman. *“Squares that look round: transforming spherical images”*, Cornell University Library ref: [arXiv:1605.01396v](https://arxiv.org/abs/1605.01396v) Mai-2016.
- [4] Leonardo Souto, Leonardo Sacht e Luiz Velho. *“Moebius Transformations Applied to Omnidirectional Images”*, Technical report TR-17-02, VISGARF Instituto de Matemática Pura e Aplicada, Feb-2017.
- [5] Marcos Rey. *“O roteirista profissional: televisão e cinema”*. 3. ed. São Paulo: Ática, 1997.
- [6] Wooksang Chang. *“Virtual Reality Filmmaking Methodology (Animation Producing)”*. TechArt: Journal of Arts and Imaging Science, Vol. 3, No. 3, Ago 2016.
- [7] Austin Baur. *“Exploring Cinematic VR: an analysis of the tools, processes, and storytelling techniques of Virtual Reality”*. Thesis, College of Motion Picture Arts, Florida State Univ., 2016.
- [8] Vanessa C. Pope, Robert Dawes, Florian Schweiger, Alia Sheikh. *“The Geometry of storytelling: Theatrical Use of Space for 360-degree Videos and Virtual Reality”*. CHI 2017, Denver, CO, USA. Mai-2017.
- [9] Oriyomi Adewale Adenuga. *“Adapting Cinematic and Theatrical Lighting to Virtual Reality”*. Thesis MsC, Texas A7M University, Dez-2016.
- [10] Ana Serrano, Vincent Sitzmann, Jaime Ruiz-Borau, Gordon Wetzstein, Diego Gutierrez, Belen Masia. *“Movie Editing and Cognitive Event Segmentation in Virtual Reality Video”*. ACM Transactions on Graphics, vol. 36, No. 4, Article 47, Jul-2017.
- [11] Rorik Henrikson, Bruno De Araujo, Fannu Chevalier, Karan Singh, Ravin Balakrishmann. *“Multi-Device Storyboards for Cinematic Narratives in VR”*. ACM SIGCHI, 2016.
- [12] Mirjam Vosmeer, Ben Schouten. *“Project Orpheus A Research Study into 360o Cinematic VR”*. TVX '17 Proceedings of the 2017 ACM International Conference on Interactive Experiences for TV and Online Video. Jun-2017.