

# Interpolação de Vistas Baseada no Rastreamento de Partículas

## Introdução

Este trabalho tem como objetivo apresentar o estudo de um novo método para realizar interpolação de vistas. A ideia central é rastrear partículas ao longo das diferentes vistas, usando um rastreador baseado em [1], realizar uma triangulação da imagem com base nos pontos rastreados, e fazer o *warping* da textura dos triângulos ao longo das diferentes vistas.

## Motivação

Filmando um determinada cena com um *array* de câmeras como no esquema da Fig. 1, obtém-se imagens de diferentes pontos de vista. Nosso interesse é, de posse de imagens da cena real, poder estimar como seria a imagem obtida com uma câmera em qualquer ponto intermediário às câmeras utilizadas. As possibilidades de aplicação disto são variadas, podemos, por exemplo, exibir um vídeo filmado com esta configuração permitindo que o espectador decida de que ponto no espaço deseja assistir, inserindo uma nova dimensão ao vídeo, como se o espectador realmente estivesse no local, podendo movimentar-se pela cena durante a exibição. Estas imagens estimadas entre duas imagens reais são o que chamamos de vistas interpoladas.

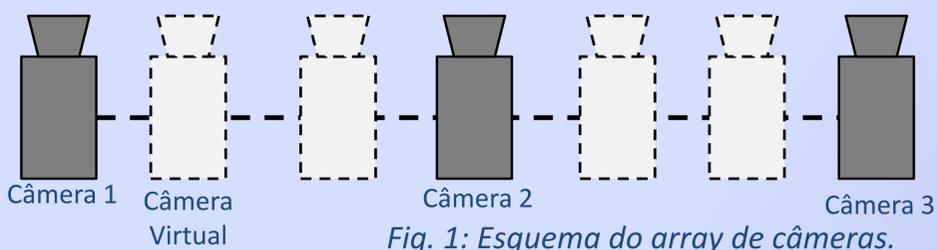


Fig. 1: Esquema do array de câmeras.

## Rastreamento de Partículas

Como mencionado, a base deste método para gerar vistas interpoladas é a tese proposta em [1], e o código fonte disponibilizado pelos autores. Brevemente, o que o artigo propõe é uma abordagem nova para a estimativa de movimento em vídeos. A ideia principal é exatamente identificar partículas, ou pontos, que são comuns entre os vários *frames* de um vídeo, no nosso caso, entre imagens obtidas de pontos diferentes da mesma cena estática, mas o princípio é o mesmo.



Fig. 2: Partículas na imagem.



Fig. 3: Identificação das bordas

Foram feitas algumas alterações no código fonte do rastreador de partículas, para que o conjunto de partículas inicial coincida com as posições das bordas da imagem, ilustrado na Fig. 3, e também são ajustadas as posições das partículas por uma reta, devido ao espaçamento constante entre as câmeras.

## Gerando Vistas Sintéticas

Para gerar as vistas sintéticas, já de posse da informação sobre as correspondências entre as imagens, realizamos uma triangulação de Delaunay para segmentar a imagem em triângulos cujos vértices são os pontos rastreados. Então, tendo a informação de posição de um determinado triângulo na imagem de uma câmera, e sabendo a posição deste na imagem seguinte, simplesmente interpolamos estas posições para obter a reta pela qual a posição de cada vértice passa, como na Fig. 4. Então utilizamos mapeamento de texturas com *OpenGL* para *renderizar* cada triângulo da imagem original na sua posição calculada nas vistas virtuais.

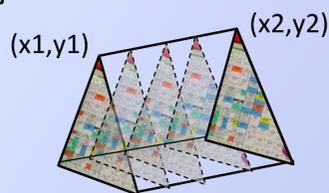


Fig. 4: Renderizando os triângulos.

Para aumentar a precisão, este processo é realizado em dois sentidos, da imagem da primeira câmera até a última, e da última para a primeira, e então faz-se uma soma ponderada das duas imagens sintéticas, a cada par de *frames* de modo que dado um determinado par, se a vista a ser gerada está mais próxima da imagem real à sua esquerda, a imagem calculada no sentido esquerda-direita possui um peso maior, e o inverso no caso da direita.

## Conclusões

Este processo ainda precisa de vários aperfeiçoamentos, como é possível nos resultados exibidos na Fig. 5. O problema mais aparente é a falta de partículas nas extremidades da imagem, que causa estas partes escuras e "serrilhadas", pois faltam partículas logo faltam triângulos. No entanto, os resultados são promissores, pois



Fig. 5: Resultados para imagens ideais (imagens de duas câmeras reais com vista virtual entre as duas sintetizada no centro).

ao contrário de outras técnicas utilizadas para a interpolação de vistas, com este método não é necessário o cálculo de mapas de disparidade denso, entre outras abordagens da Visão Estéreo, o que torna o processo muito mais simples.

## Referências

- [1] Peter Sand and Seth Teller, *Particle Video: Long-Range Motion Estimation using Point Trajectories*, CVPR 2006.
- [2] D. Scharstein and R. Szeliski. *High-accuracy stereo depth maps using structured light*. CVPR 2003

## Websites:

- rvsn.csail.mit.edu/pv/-vision.middlebury.edu/stereo/data/