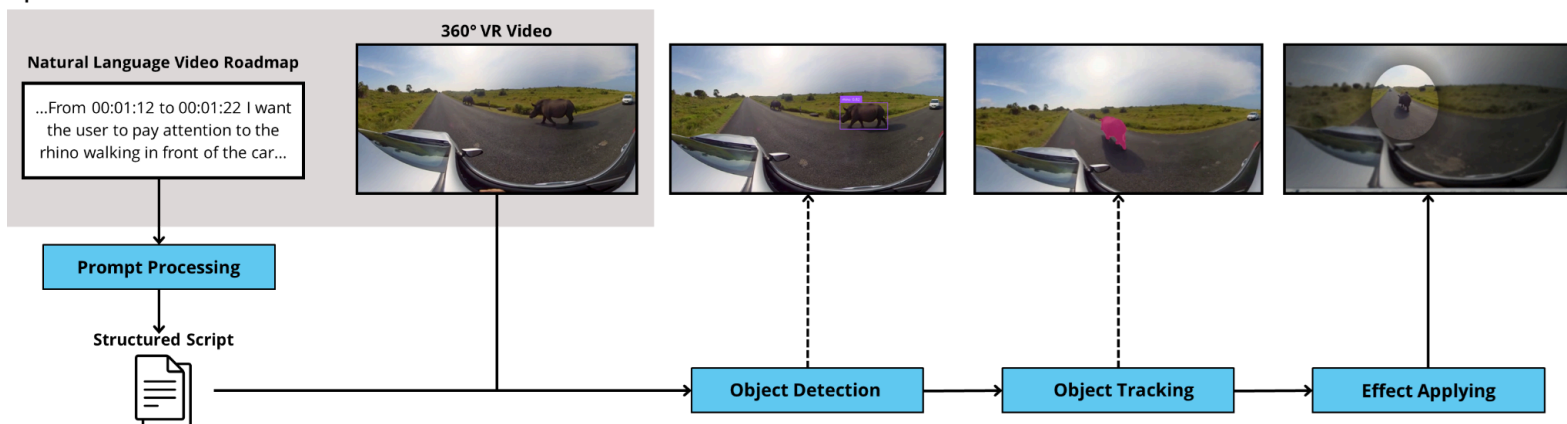


Pipeline Voltado para a Tarefa de Guiar a Atenção do Usuário em Vídeos em Realidade Virtual

Input



1º Passo) Usuário - Elencar, no vídeo, quais objetos serão alvos durante a reprodução. Para cada objeto deve ser passado o intervalo de tempo no qual ele deve ser focado.

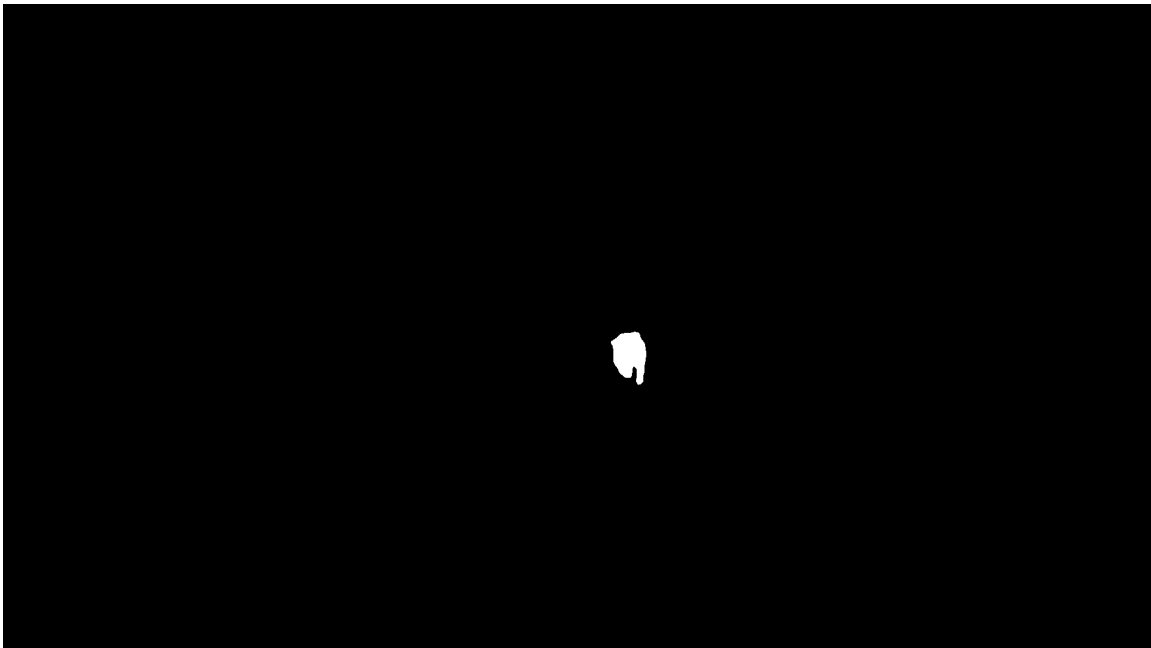
Ex: *De 00:01:00 até 00:01:15 eu quero que o rinoceronte que está andando na frente do carro seja destacado. Depois, de 00:02:13 até 00:03:00 eu quero que a girafa no meio das árvores seja destacada*

2º Passo) Manipulação de Prompt (LLM) - A descrição fornecida pelo usuário é passada para um LLM que irá extrair as informações necessárias (objetos e intervalos de tempo) e estruturá-las no formato necessário para o Pipeline, como no exemplo abaixo:

from	to	object_description
00:00:26	00:00:33	the rhino in front of the car
00:00:41	00:00:48	the rhino behind the car
00:01:12	00:01:19	the deer
00:01:33	00:01:40	the giraffe walking in jungle
00:02:16	00:02:23	the animal between the trees
00:02:34	00:02:41	the buffaloes
00:03:03	00:03:10	the turtle

3º Passo) Identificação Inicial (Detecção de Objeto) - Para cada objeto alvo, o primeiro frame do intervalo fornecido é passado para um modelo multimodal, que recebe o input com a descrição do objeto e o frame e realiza a detecção do objeto alvo na imagem, fornecendo sua bounding box.

4º Passo) Tracking do Objeto (Segmentação/Tracking em Vídeo) - É extraído o ponto central da bounding box fornecida na etapa anterior e este ponto serve como input para o modelo que irá realizar o tracking do objeto alvo durante o intervalo solicitado, mantendo-o no foco do efeito que será aplicado posteriormente para guiar a atenção do usuário:



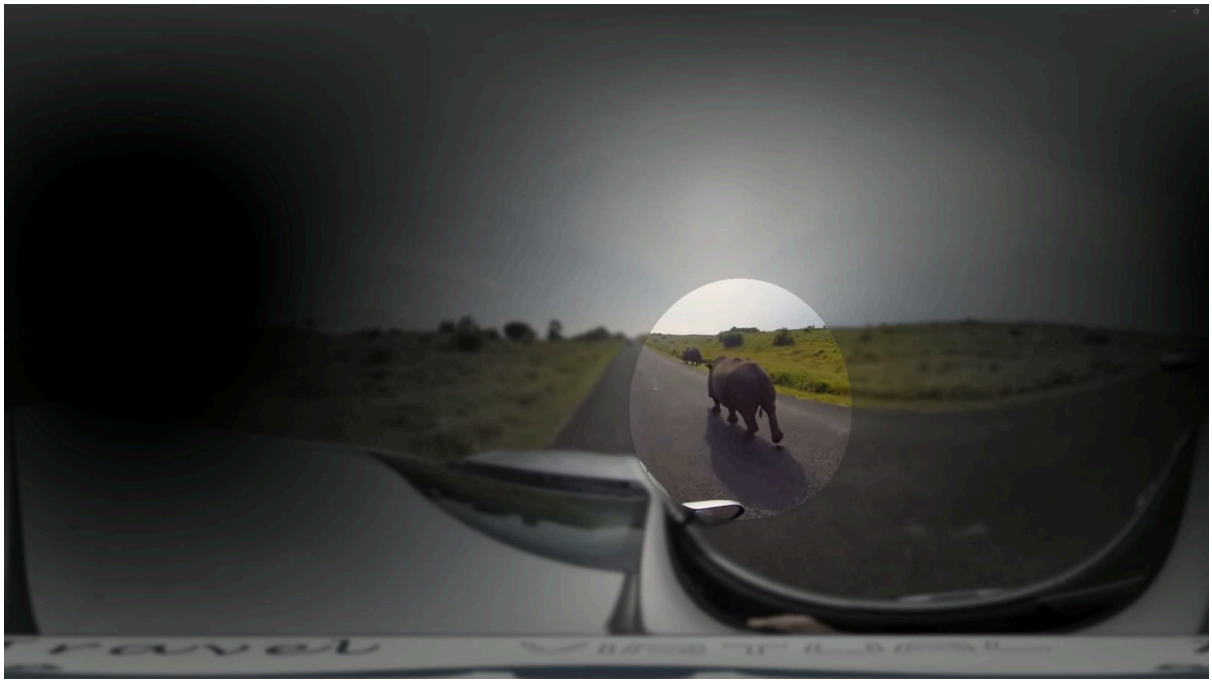
Máscara de segmentação produzida



Máscara de segmentação sobreposta no frame original

Vídeo com o resultado de um exemplo: [Vídeo](#)

5º Passo) Aplicação dos Efeitos (Processamento Digital de Imagens) - Para cada frame processado é aplicado um conjunto de efeitos para destacar o objeto desejado, os efeitos são: desfoque constante, gradiente de cinza, gradiente de escurecimento e destaque radial. Uma pequena área radial ao redor do objeto é preservada para mantê-lo em destaque com relação ao restante do vídeo. É feito um aumento gradual da intensidade dos efeitos tanto no início quanto no fim para que não haja mudanças abruptas. Exemplo de um resultado:



Link para pasta com mais exemplos: [Pasta de vídeos](#)