

Interpolação de Vídeos

Conrado Santos Boeira

Pontifícia Universidade Católica do Rio Grande do Sul (PUCRS)

Porto Alegre, Brasil

conrado.boeira@edu.pucrs.br

Resumo—Existem várias maneiras de se melhorar a qualidade de um vídeo. Uma delas é o aumento de taxa de frames através da interpolação de frames consecutivos. Nesse trabalho são apresentadas duas maneiras diferentes de interpolação. A primeira consiste em interpolar a cor de cada pixel entre dois frames. A segunda avalia possíveis movimentos de blocos de pixels e gera uma interpolação baseada nesse movimento.

I. INTRODUÇÃO

Uma das maneiras mais comuns de entretenimento e de distribuição de informações importantes é através de vídeos. Dessa maneira, qualquer método capaz de aumentar a qualidade deles traz um grande benefício. Assim, nesse trabalho serão apresentados duas possíveis maneiras de melhorar um aspecto específico de vídeos, a taxa de frames por segundo.

Para isso, será usado o conceito de interpolação. Interpolação é um tipo de método que estima pontos novos dentro de um intervalo a partir de pontos já tabelados deste intervalo. A interpolação será usada nas duas soluções, porém de duas maneiras diferentes. A primeira consiste em aplicar uma interpolação linear entre a cor de cada pixel em 2 frames contíguos. Já a segunda solução é baseada no trabalho proposto por Wang et al. [1] e procura estimar o deslocamento de blocos de pixels entre os dois frames que se deseja interpolar.

Neste trabalho, os dois métodos desenvolvidos são apresentados na Secção 2. Os resultados dos dois métodos incluindo exemplos dos frames obtidos com eles são exibidos na Secção 3. Finalmente, conclusões finais e trabalhos futuros são discutidos na Secção 4.

II. SOLUÇÕES PROPOSTAS

O primeiro método proposto foi desenvolvido individualmente. O segundo foi uma contribuição entre os alunos da turma. As duas soluções foram desenvolvidas usando a linguagem Python3 e com o auxílio da biblioteca para cálculos científicos Numpy.

A. Primeira Solução

O primeiro método desenvolvido foi o mais simples entre os dois. Nessa solução é feita a iteração sobre todos os frames do vídeo escolhido e, a cada par de frames, é inserido um frame intermediário. Cada pixel desse frame tem um valor que lhe é atribuído a partir da definição do ponto médio entre o valor dos pixels de mesma posição dos frames vizinhos. Dessa maneira cada pixel do frame i a ser criado entre os frames j e $j + 1$ pode ser definido como:

$$P_{i,x,y} = \frac{P_{j,x,y}}{2} + \frac{P_{j+1,x,y}}{2}$$

Onde $P_{i,x,y}$ define o valor do pixel do frame i na posição (x, y) .

B. Segunda Solução

A segunda solução consiste em procurar interpolar o movimento de blocos de pixels entre os dois frames que se quer interpolar. Primeiramente se deseja separar o primeiro dos frames em blocos de tamanho arbitrário t . Assim, para cada bloco $B_{f,x,y,t}$ pertencente ao frame f e que tem o canto superior esquerdo localizado em (x, y) e tamanho t calculamos a soma dos valores dos pixels pertencente ao bloco. Essa soma pode ser descrita como:

$$SOMA(B_{f,x,y,t}) = \sum_{i=x}^{x+t} \sum_{j=y}^{y+t} P_{f,i,j}$$

Feito isso, é feita uma procura no frame $f + 1$ pelo bloco de mesmo tamanho e tal que

$$\begin{aligned} &\|SOMA(B_{f,x,y,t}) - SOMA(B_{f+1,x',y',t})\| \leq \\ &\|SOMA(B_{f,x,y,t}) - SOMA(B_{f+1,i,j,t})\| \end{aligned}$$

para qualquer valor de i e j que esteja dentro dos limites de $f + 1$ e com uma distância do ponto (x, y) menor ou igual a t . Isso é feito para estimar qual foi o deslocamento do bloco entre os dois frames. Com a posição desses dois blocos, é possível calcular o vetor de movimento do bloco.

Uma vez que o vetor de movimento de cada bloco do frame f foi calculado, é possível construir o frame intermediário. Para isso, todos pixels de cada bloco são deslocados metade do deslocamento encontrado. Isso gera duas possíveis dificuldades, pode haver conflito entre dois blocos que se deslocam para mesmas posições e podem haver posições no frame que não foram ocupadas por nenhum bloco. A solução para a primeira dificuldade consiste em simplesmente sobrescrever os primeiros dados gravados em caso de conflito, já que foi constatado que essa situação não acontece comumente. O segundo problema foi solucionado preenchendo todos os pixels não ocupados com os valores encontrados usando a primeira solução proposta.

III. RESULTADOS EXPERIMENTAIS

As duas soluções foram testadas usando alguns vídeos fornecidos. É possível ver os resultados encontrados nas Figuras 1 a 6. Tanto as Figuras 1 e 4 quanto as Figuras 3 e 6 são iguais entre si e foram replicadas para melhor mostrar a qualidade do frame interpolado.



Fig. 1. Frame 1 do vídeo



Fig. 2. Frame interpolado usando a primeira solução



Fig. 3. Frame 2 do vídeo



Fig. 4. Frame 1 do vídeo



Fig. 5. Frame interpolado usando a segunda solução



Fig. 6. Frame 2 do vídeo

Na Figura 2 é possível ver o resultado da primeira solução. Devido a interpolação direta das cores dos pixels, é claramente visível que a forma da bola aparece duas vezes borrada. Isso dá um efeito ao vídeo de um certo sombreamento ao movimento da bolinha.

Na Figura 4 podemos ver o resultado da segunda solução. Comparando com as Figuras 4 e 6, é possível ver que a bola parece estar exatamente na metade do caminho entre a sua posição no frame anterior e no seguinte. Essa é justamente a vantagem do uso da interpolação pelo movimento. Os blocos aos quais a imagem da bola pertence são deslocados seguindo um mesmo vetor, proporcionando esse resultado.

Comparando os resultados é possível ver que a segunda solução gera um resultado consideravelmente melhor especialmente em vídeos com movimentos rápidos. Os frames criados através desse método não exibem o sombreamento que acontece na primeira solução criando uma imagem que parece mais natural ao espectador do vídeo.

Apesar de gerar um resultado superior, a segunda solução também não é livre de ter seus problemas. Essa solução é mais cara computacionalmente e requer um tempo muito maior para ser executada. Além disso, como é possível ver pela Figura 7, é possível que blocos com a mesma cor de um objeto sejam deslocados para uma mesma posição, o que pode levar aos efeitos vistos na imagem. Porém esse efeito é difícil de se notar por parte de quem assiste o vídeo interpolado e o resultado final acaba sendo satisfatório.

IV. CONCLUSÕES

O trabalho aqui apresentado resolveu o problema proposto de aumentar a quantidade de frames de um vídeo através do uso de técnicas de interpolação. Foram apresentadas dois



Fig. 7. Frame gerado pela solução 2

métodos para esse fim e foi constatado que o método de interpolação do movimento de blocos entre os frames obteve o melhor resultado para o problema.

Existem muitas possíveis melhorias que poderiam ter sido feitas na solução. Uma possibilidade seria o uso da estrutura Quad-Tree para definir em função de cada frame o tamanho dos blocos a serem utilizados. Isso permitiria que o tamanho dos blocos se adaptasse ao tamanho dos objetos da imagem, o que impediria a fragmentação vista na Figura 7. Outra possibilidade seria a de usar mais frames (anteriores ou posteriores) para calcular o vetor de movimento. Essas duas possibilidades foram exploradas em [2].

REFERÊNCIAS

- [1] Wang, Demin, et al. "Motion-compensated frame rate up-conversion—Part II: New algorithms for frame interpolation." *IEEE Transactions on Broadcasting* 56.2 (2010): 142-149.

- [2] Min, Kyung-Yeon, et al. "Bidirectional mesh-based frame rate up-conversion." *IEEE MultiMedia* 22.2 (2014): 36-45.