

Trabalho 3

1 Especificação do Problema

Um vídeo digital pode ser definido como uma coleção de imagens de mesmas dimensões, sendo agrupadas de acordo com uma sequência temporal. Cada uma dessas imagens é denominada quadro, o qual corresponde à menor unidade estrutural do vídeo. Os quadros podem ser agrupados em tomadas, que são sequências de quadros capturados de forma contígua e que representam uma ação contínua no tempo ou no espaço. Finalmente, um grupo de tomadas que são semanticamente correlacionadas constitui uma cena. A Figura 1 mostra a estrutura geral de um vídeo.

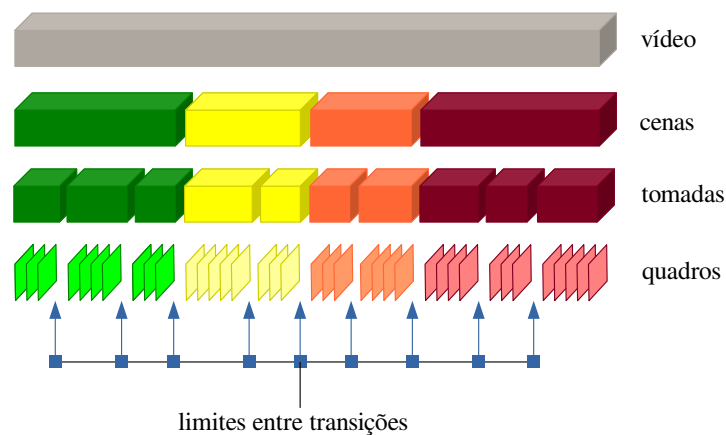


Figura 1: Estrutura geral de um vídeo.

O objetivo deste trabalho é detectar transições abruptas que ocorrem entre quadros que compõem um vídeo. Diferentes abordagens serão implementadas e testadas. Os vídeos devem ser inicialmente convertidos em escala de cinza.

Há diferentes tipos de transições em vídeos. Um corte (*cut*) é uma mudança abrupta que ocorre em um único quadro. Um desbotamento (*fade*) corresponde à mudança lenta em brilho, normalmente resultando em um quadro totalmente preto. Uma dissolução (*dissolve*) ocorre quando um quadro gradualmente desaparece, enquanto seu quadro consecutivo torna-se mais visível. Uma limpeza (*wipe*) ocorre quando pixels de um quadro são trocados com pixels do quadro adjacente conforme um padrão regular, por exemplo, uma coluna a partir da borda esquerda dos quadros.

1.1 Diferenças entre Pixels

Dois quadros consecutivos são considerados significativamente diferentes se a contagem do número de pixels que sofrem alteração em intensidade, conforme uma certa tolerância T_1 , é maior que um segundo limiar T_2 .

1.2 Diferenças entre Blocos

Dois quadros consecutivos são divididos em 8×8 ou 16×16 blocos sem sobreposição. O erro quadrático da diferença de dois blocos correspondentes entre as imagens deve ser calculado e testado contra um limiar T_1 . Se o erro for menor do que o limiar, então os quadros são considerados similares; caso contrário, eles são considerados distintos. Um segundo limiar T_2 deve ser definido e utilizado para avaliar o número de blocos que possuem erro quadrático maior do que T_1 . Se esse número exceder T_2 , então o quadro será considerado uma transição abrupta.

1.3 Diferenças entre Histogramas

A diferença D_i dos histogramas de intensidades para dois quadros consecutivos i e $i + 1$ do vídeo pode ser expressa como

$$D_i = \sum_{j=1}^B |H_i(j) - H_{i+1}(j)|$$

em que B denota o número total de *bins* no histograma e $H_i(j)$ é o valor do histograma para o i -ésimo quadro no nível j .

Um limiar T deve ser utilizado para determinar se a diferença dos histogramas de intensidade D_i indica uma transição abrupta no vídeo. O limiar pode ser estimado como

$$T = \mu + \alpha\sigma$$

em que μ e σ são, respectivamente, o valor médio e o desvio padrão da diferença dos histogramas de intensidade. O valor de α tipicamente varia de 3 a 6.

1.4 Diferenças entre Mapas de Bordas

Diferenças entre os mapas de bordas podem ser utilizadas para verificar alterações entre quadros dos vídeos. A relação entre o número de pixels de bordas entre dois quadros consecutivos i e $i + 1$ é calculada. Um limiar T deve ser utilizado para determinar se a diferença dos mapas de bordas indica uma transição abrupta no vídeo.

2 Resultados

O gráfico mostrado na Figura 2 ilustra os valores obtidos pela métrica calculada para cada quadro do vídeo. As transições são representadas pelos picos representativos no gráfico.

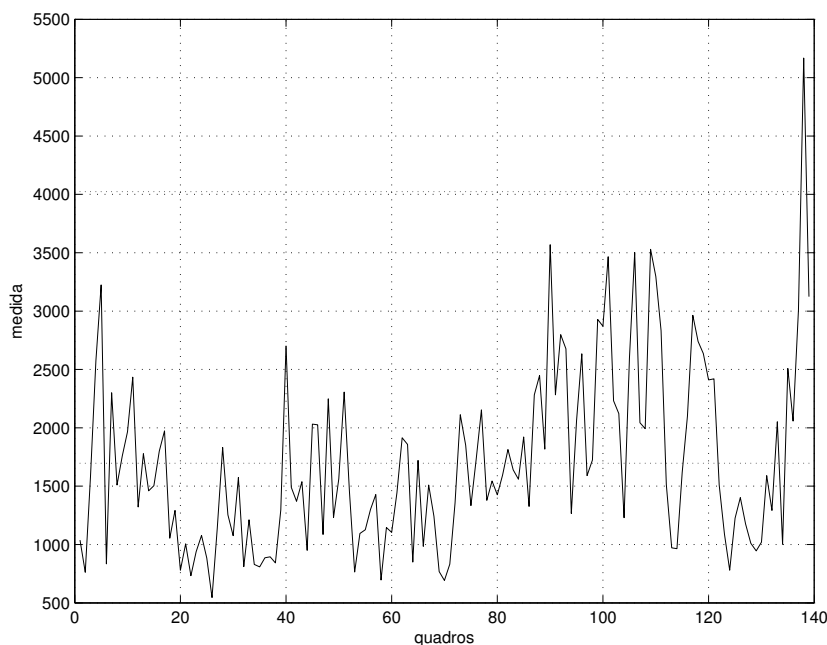


Figura 2: Resultado da métrica para cada quadro do vídeo.

3 Entrada de Dados

Os vídeos de entrada estão no formato MP4 ou MPEG Layer 4 (*Moving Picture Experts Group*). Alguns exemplos encontram-se disponíveis no diretório: http://www.ic.unicamp.br/~helio/videos_mp4/

4 Saída de Dados

Os vídeos de saída contendo os quadros de transição obtidos devem estar no formato MP4 ou MPEG Layer 4 (*Moving Picture Experts Group*). Os gráficos gerados devem ser exibidos na tela.

5 Especificação da Entrega

- A entrega do trabalho deve conter os seguintes itens:
 - código fonte: o arquivo final deve estar no formato *zip* ou no formato *tgz*, contendo todos os programas necessários para sua execução.
 - relatório: deve conter uma descrição dos algoritmos e das estruturas de dados, considerações adotadas na solução do problema, testes executados, eventuais limitações ou situações especiais não tratadas pelo programa.
- O trabalho deve ser submetido por meio da plataforma *Google Classroom*.
- Data de entrega: 08/06/2022

6 Observações Gerais

- Os programas serão executados em ambiente Linux. Os formatos de entrada e saída dos dados devem ser rigorosamente respeitados pelo programa, conforme definidos anteriormente. Não serão aceitos trabalhos após 5 dias da data de entrega.
- Os seguintes aspectos serão considerados na avaliação: funcionamento da implementação, clareza do código, qualidade do relatório técnico.