Segmentação Temporal de Vídeos

Maria Clara Alvarenga Gouveia*, Helio Pedrini

Resumo

A segmentação temporal de vídeos é o processo automático de identificação de transições abruptas e graduais presentes entre as unidades estruturais do vídeo e constitui uma técnica fundamental para auxiliar a atividade de análise de vídeos. Propõe-se, neste projeto, a investigação das principais estratégias existentes para segmentação temporal de vídeos, bem como o desenvolvimento de uma nova metodologia para detectar as principais transições que ocorrem em sequências de vídeos.

Palavras-chave:

Segmentação de vídeos, transições graduais, cortes abruptos

Introdução

O principal objetivo deste trabalho é investigar e desenvolver técnicas de segmentação temporal de vídeos digitais. Experimentos foram realizados em conteúdos de diferentes gêneros, tais como documentários, filmes, noticiários, esportes, entre outros.

O método proposto emprega diversas estratégias para identificar transições entre tomadas a partir da extração de características dos quadros, tais como diferença entre pixels correspondentes em dois quadros, diferença entre histogramas, variações de bordas de objetos, alterações em características medidas em blocos dos quadros, entre outras.

Material e Métodos

A implementação dos códigos foi realizada na linguagem de programação Python e foram utilizados os pacotes SciPy, NumPy, SciKit-Image, SciKit-Learn e OpenCV.

Neste trabalho, um vasto levantamento de estudos bibliográficos foi efetuado, bem como testes e experimentos com diferentes métodos. O método de diferenças entre pixels foi o mais empregado em nossos experimentos. Neste algoritmo, duas tomadas adjacentes são consideradas diferentes se a percentagem de pixels correspondentes com diferença absoluta (T_{diff} , em relação a um limiar T_1) for superior a um limiar T_2 .

Resultados e Discussão

O algoritmo utilizado foi capaz de identificar todos ou a maioria dos cortes abruptos. A Figura 1 ilustra um exemplo, em que os pontos vermelhos estão acima do limiar $T_1 = 80$.

Apesar disso, percebeu-se que as transições graduais muitas vezes não foram detectadas ou, quando o limiar T_2 é reduzido, vários falsos positivos foram detectados (como ilustrado na Figura 2).

Até o momento, utilizou-se um limiar T_2 global, ou seja, fixo para todo o vídeo.

Figura 1. Gráfico de T_{diff} em um vídeo com cortes abruptos ($T_2 = 20$).

Figura 2. Gráfico de T_{diff} em um vídeo com transição gradual ($T_2 = 20$).

Conclusões

A partir dos experimentos, constatou-se que, embora o método funcione bem para detecção de cortes abruptos, ele não tem um comportamento adequado com relação a transições graduais, como *fade* e *dissolve*, utilizando um limiar T_2 global. Como trabalho futuro, sugere-se a aplicação de um limiar local (adaptativo para diferentes blocos do vídeo) a diferentes métodos de segmentação.

Agradecimentos

Agradeço ao CNPq pelo financiamento da bolsa de Iniciação Científica.

¹ K. N. Ngan and H. Li. Video Segmentation and its Applications. Springer Science & Business. Media, 2011.

² M. V. M. Cirne and H. Pedrini. Summarization of Videos by Image Quality Assessment. In Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, pages 901–908. Springer, 2014.