

Processamento e Análise de Imagens (MC940)

Análise de Imagens (MO445)

Universidade Estadual de Campinas

Instituto de Computação

Professor: Helio Pedrini

Aluno: Renan Gomes Pereira 103927

Trabalho 2

Introdução

O objetivo deste trabalho é detectar transições abruptas que ocorrem entre quadros que compõem um vídeo. Foram implementados 3 métodos diferentes que calculam o número de transições para os vídeos *toy* e *xylophone*. Os métodos implementados são: Diferença entre pixels; diferenças estatísticas; e histogramas.

Procedimentos

Foi criado um programa utilizando o software MATLAB R2014a para o sistema operacional Windows 8.1 64 bits. Este programa permite que o usuário escolha se ele deseja que o programa plote os gráficos convenientes e quais vídeos analisar. Essas opções estão disponíveis no *script main.m* através de variáveis booleanas.

O programa está dividido em:

main.m: coordena a chamada das funções e a visualização do resultado para ambos os vídeos.

read_video_conv_rgb2gray(path): lê um vídeo RGB cujo o endereço é *path* e retorna uma matriz com os quadros do vídeo convertidos para níveis de cinza.

pixel_dif(frames, t1, t2, plot_images): dados os limiares *t1* e *t2*, retorna o número de transições entre dois quadros consecutivos pelo método da diferença entre pixels (explicado nas seções abaixo). Se *plot_images* é true, plota o gráfico do número de *pixels* com alteração de intensidade acima do limiar *t1*.

statistic_dif(frames, bs, t1, t2, plot_images): dados os limiares $t1$ e $t2$, retorna o número de transições entre dois quadros consecutivos pelo método das diferenças estatísticas (explicado nas seções abaixo).

hist_dif(frames, alfa, nbins, plot_images): dado o valor de $alfa$ e um vetor com o número de $nbins$ que deseja calcular os histogramas, retorna o número de transições entre dois quadros consecutivos pelo método da diferença dos histogramas de intensidade (explicado nas seções abaixo).

Diferença entre pixels

Foi implementada uma função para calcular o número de transições baseado na diferença entre pixels entre dois quadros consecutivos conforme especificado em [1]. Para ambas as imagens, *toy* e *xylophone*, os valores dos limiares $t1$ e $t2$ foram ajustados através da análise visual dos resultados de diversos testes até que se chegasse em um valor que foi julgado adequado. É importante ressaltar que o processo de obtenção de $t1$ e $t2$ pode ser feito com o uso outros métodos como aprendizado por redes neurais, porém isso foge do escopo deste trabalho.

Os resultados obtidos foram:

Vídeo: <i>Toy</i>
$t1 = 50$
$t2 = 500$
Resultado para a diferença entre pixels: 47
Gráfico: Figura 1.

Vídeo: <i>Xylophone</i>
t1 = 50
t2 = 400
Resultado para a diferença entre pixels: 81
Gráfico: Figura 2.

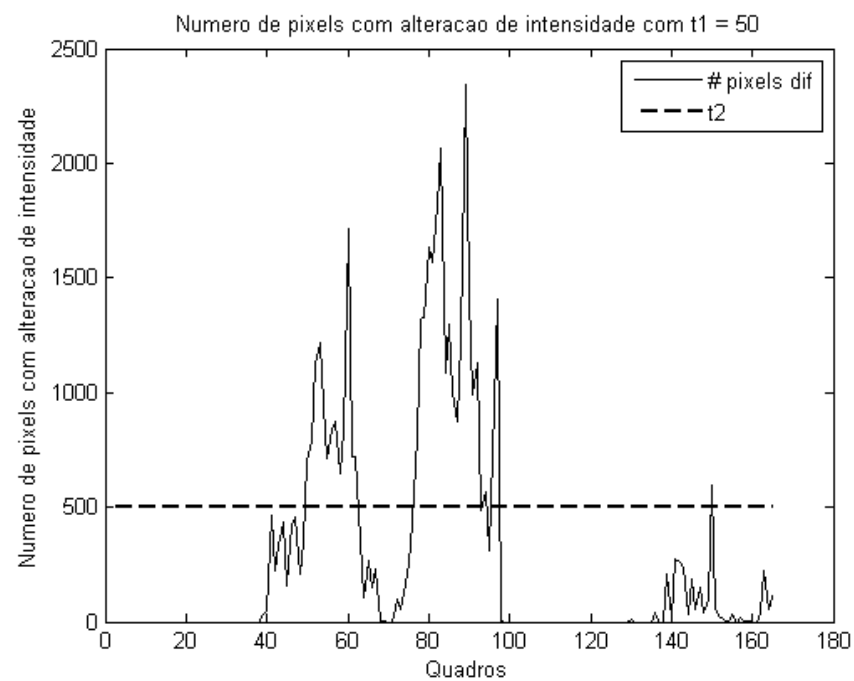


Figura 1: Gráfico do número de pixels com alteração de intensidade maior que t1 para o vídeo toy.

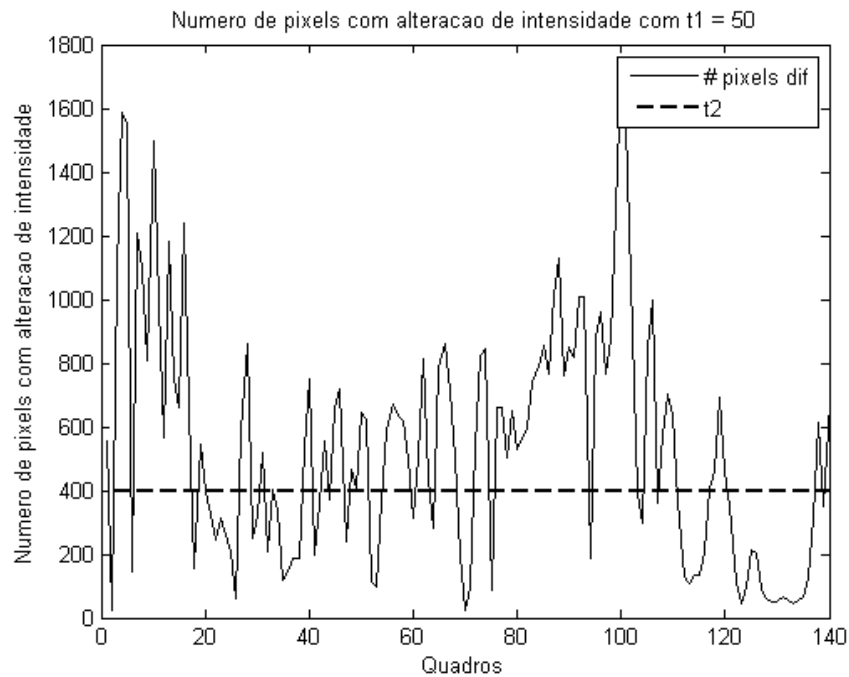


Figura 2: Gráfico do número de pixels com alteração de intensidade maior que t_1 para o vídeo xylophone.

Como era esperado após assistir ambos os vídeos, o número de transições para o vídeo xylophone foi superior ao resultado obtido para o vídeo toy.

A Figura 1 mostra uma característica importante do vídeo *toy* que é o fato do brinquedo não se mover nos instantes iniciais e finais do vídeo (apesar da câmera tremer no final do vídeo). Por outro lado, a Figura 2 indica que houve movimento durante todo o vídeo do *xylophone*.

Diferenças estatísticas

Foi implementada uma função para calcular o número de transições baseado nas diferenças estatísticas conforme especificado em [1]. Cada quadro foi dividido em blocos de tamanho 8x8. Assim como no caso anterior da diferença entre *pixels*, os valores dos limiares t_1 e t_2 foram calculados após a análise de diversos testes e os valores escolhidos foram aqueles que apresentaram melhores resultados.

Os resultados obtidos foram:

Vídeo: <i>Toy</i>
$t1 = 100$
$t2 = 120$
Resultado para as diferenças estatísticas: 56
Gráfico: Figura 3.

Vídeo: <i>Xylophone</i>
$t1 = 100$
$t2 = 70$
Resultado para as diferenças estatísticas: 77
Gráfico: Figura 4.

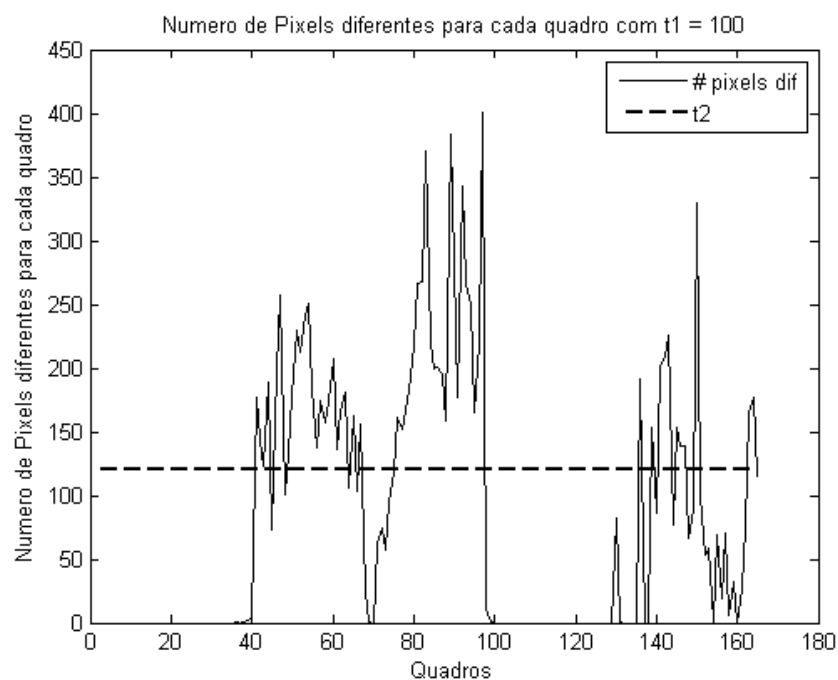


Figura 3: Gráfico do número de *pixels* diferentes para cada quadro acima do limiar $t1$ para o vídeo *toy*.

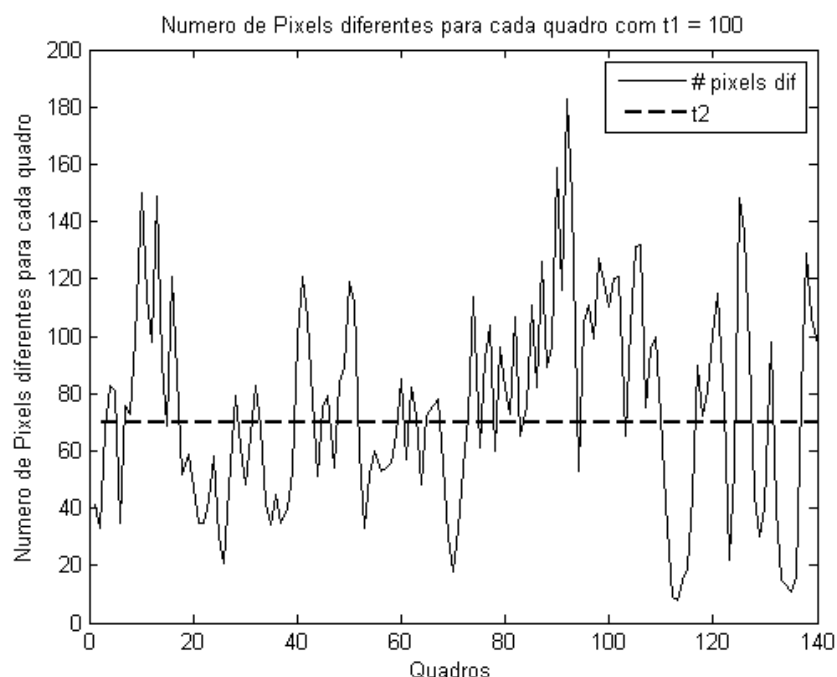


Figura 4: Gráfico do número de *pixels* diferentes para cada quadro acima do limiar t_1 para o vídeo *xylophone*.

Os resultados foram parecidos com o método anterior da diferença entre *pixels*, pode-se notar a semelhança na forma entre os gráficos mostrados na Figura 1 com a Figura 3 e na Figura 2 com a Figura 4. Novamente, como esperado, o número de transições calculadas para o vídeo *xylophone* foi maior que o número de transições calculadas para o vídeo *toy*.

Histogramas

Foi implementada uma função para calcular o número de transições baseado na diferença dos histogramas de intensidade para dois quadros consecutivos conforme especificado em [1]. Para este caso, o limiar depende do valor médio e do desvio padrão da diferença dos histogramas de intensidade, portanto, o limiar não foi escolhido de forma totalmente arbitrária. Contudo, o valor do limiar para este método ainda depende de um valor α escolhido pelo usuário. Assim como nos casos anteriores, o valor de α escolhido foi através da análise dos resultados de diversos testes.

Para calcular a diferença dos histogramas de intensidade para dois quadros consecutivos, deve-se calcular o histograma dos quadros i , $i+1$ em uma iteração. Na próxima iteração, calcula-se novamente os histogramas dos novos quadros com valor de i incrementado de uma unidade. Porém um dos quadros já teve seu histograma calculado na iteração passada. Uma simples otimização nessa parte é calcular o histograma do primeiro quadro fora do *loop*; a faixa do *loop* é de k variando do segundo quadro até o último; para uma iteração i , temos h_{i-1} calculado da iteração passada, portanto, devemos calcular apenas h_i ; no final da iteração $h-1 = h_i$.

O número de transições obtidas nesta parte também depende do número de *bins* dos histogramas. Com o objetivo de testar a diferença nos resultados devido a alteração do número de *bins*, a função implementada recebe um vetor com os números de *bins* desejados e retorna um outro vetor com o número de transições obtidas para cada número de *bins* do vetor de entrada. Para não tornar o relatório muito extenso, os gráficos mostrados abaixo são apenas para os histogramas de 256 *bins*.

Os resultados obtidos foram:

Vídeo: <i>Toy</i>
Alfa = 1
Resultado para histogramas com 4 bins: 23 Resultado para histogramas com 8 bins: 28 Resultado para histogramas com 32 bins: 33 Resultado para histogramas com 256 bins: 33
Gráfico: Figura 5.

Vídeo: <i>Xylophone</i>
Alfa = 0,4
<p>Resultado para histogramas com 4 bins: 46</p> <p>Resultado para histogramas com 8 bins: 45</p> <p>Resultado para histogramas com 32 bins: 66</p> <p>Resultado para histogramas com 256 bins: 59</p>
Gráfico: Figura 6.

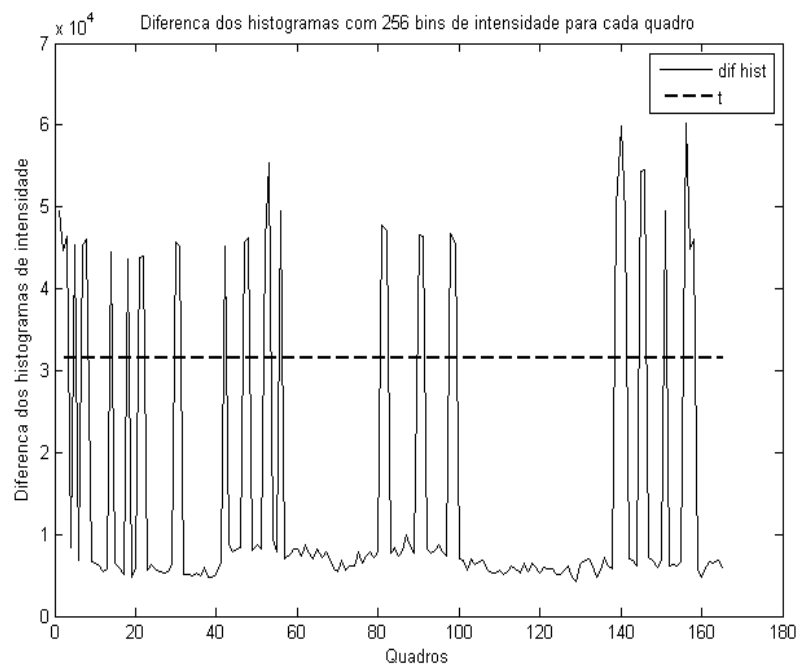


Figura 5: Diferenças dos histogramas com 256 *bins* para o vídeo toy.

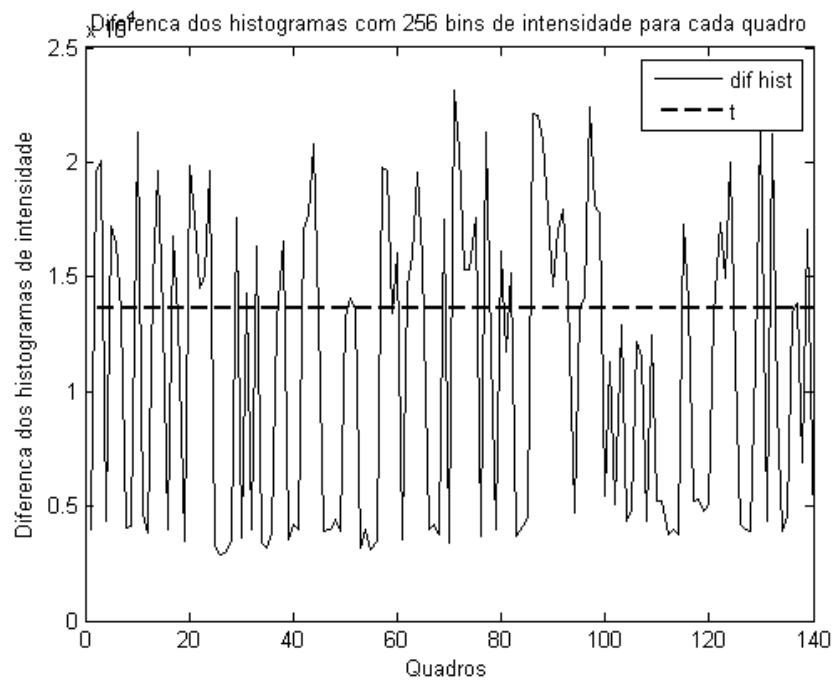


Figura 6: Diferenças dos histogramas com 256 *bins* para o vídeo *xylophone*.

Para ambos os vídeos, o número de *bins* influenciou no número de transições obtidas. A relação entre esses dois valores **não** é linear diferente da intuição de que quanto maior o número de *bins*, maior o número de transições. Para que fosse possível obter um número de transições para esses vídeos, o valor de alfa foi abaixo da faixa de valores típicos de alfa que varia de 3 a 6[1]. Isso reforça a necessidade de uma análise dos resultados deste método feita por um ser humano para estimar um valor adequado para alfa.

Conclusões

O objetivo deste trabalho era detectar transições abruptas que ocorrem entre quadros que compõem um vídeo. Para isso, foram implementados 3 métodos diferentes que calculam o número de transições para os vídeos *toy* e *xylophone*. Os métodos implementados foram: Diferença entre pixels; diferenças estatísticas; e histogramas.

O MATLAB foi uma ferramenta muito poderosa para a realização deste trabalho que implementa métodos em vídeos. Apesar do fato dos vídeos considerados serem curtos, com menos de 6 segundos, o consumo de memória e tempo de processamento são relativamente altos comparados com diversas operações em imagens. Portanto, o uso de uma ferramenta que manipula matrizes de forma eficiente é de muita importância.

A maior dificuldade foi estimar os valores dos limiares t_1 e t_2 para os casos da diferença entre pixel e das diferenças estatísticas. As estimativas foram feitas através da análise dos resultados de diversos testes com limiares diferentes. Esse processo poderia ser feito com técnicas computacionais como redes neurais, porém isso foge do escopo deste trabalho.

Bibliografia

[1] Enunciado do trabalho 2. Site professor Hélio acessado em (outubro 2015)
<http://www.ic.unicamp.br/~helio/disciplinas/MC940/trabalho2.pdf>