

MC920 - Introdução ao Processamento de Imagens

Trabalho 2: Transições abruptas em vídeos.

FRANCISCO RAFAEL CARNEIRO

RA 157888

UNICAMP
2017

1. Introdução

Neste trabalho, teve-se por objetivo o estudo e análise de quatro métodos para que pudesse serem encontrados os frames que consistiam em transições abruptas entre si. Ou seja, achar cortes no vídeo, no qual ocorre uma mudança drástica da que se está observando na passagem de frames, como por exemplo, uma mudança de cena.

Os métodos utilizados foram: Diferença entre pixels, Diferença entre blocos, Diferença entre histogramas e por fim a diferença entre mapas de borda. Após calculados, gerou-se gráficos que identificam essas transições juntamente com um vídeo contendo os frames dos cortes.

2. Informações adicionais

2.1. Bibliotecas

A linguagem utilizada foi Python, na versão 2.7.13. As bibliotecas utilizadas foram:

Numpy: Manipulação de matrizes, vetores e dados estatísticos.

Matplotlib: Geração de gráficos e imagens.

Math: Operações matemáticas.

OpenCV2: Manipulação de imagens e vídeos.

Skimage: Manipulação de imagens e vídeos.

Skvideo: Leitura e escrita de vídeos.

Sobre a biblioteca Skvideo, foi utilizada a função **skvideo.io.read** para a leitura de um vídeo, onde esta função retorna o um vetor do tipo **numpy** que gera maior facilidade na manipulação. Após isso, houve a conversão de todo o vídeo para escala de cinza. E para salvar os frames definidos como transições abruptas, utilizou-se a função **skvideo.io.vwrite**.

2.2. Eventuais problemas

No processo para salvar os vídeos, ocorreram problemas para a utilização da biblioteca **OpenCV2**, pois não conseguiu limitar o número de frames por segundo. Por isso, os vídeos foram salvos utilizando a biblioteca **Skvideo** e não foi possível essa mudança, mantendo os frames por segundo originais.

Outro fator é que essa biblioteca faz a utilização do software **FFMPEG**, no qual ocorreram problemas de instalação e de utilização, porém se mostrou estável na versão 3.2.4.

3. Metodologia e Desenvolvimento

Nesta seção será discutido os métodos utilizados para o cálculo dos cortes no vídeo e também os algoritmos envolvidos.

3.1. Diferença de Pixels.

A diferença de pixels consiste no processo pelo qual são contados a quantidade de pixels que mudaram de acordo com um limiar T_1 , entre um frame e o seu subsequente. Ou seja, para cada frame e o seu subsequente, analisa-se a diferença entre um pixel $p_{k(i,j)}$ com $p_{k+1(i,j)}$, caso essa diferença seja maior que T_1 , é considerado uma mudança. Por fim, testa-se o resultado de cada frame com um limiar T_2 , caso seja maior irá indicar uma transição abrupta de frames. O algoritmo utilizado para esse problema foi:

```
def pixels_t1(video, T):  
    frames = video.shape[0]  
    vect = np.zeros(frames - 1)  
    for k in range (frames - 1):  
        vect[k] = np.count_nonzero((abs(video[k] - video[k + 1]) > T))  
  
    return vect
```

Figura 1: Algoritmo para o cálculo da diferença de pixels

3.2. Diferença de Blocos

Um frame e o seu subsequente são divididos tal que ambos fiquem com 64 blocos (8x8) ou com 256 (16x16). Após a divisão foi calculada o erro quadrático da diferença entre dois blocos de mesma posição em frames diferentes. Assim, é testado com um limiar T_1 . Caso o erro seja menor que o limiar, não é contabilizado (frames similares), caso contrário não são similares. O algoritmo abaixo representa esse cálculo:

```

def square_sum_diff(video, div, T):
    frames = video.shape[0]
    m = video.shape[1]
    n = video.shape[2]
    xbloco = m/div
    ybloco = n/div
    count = 0
    vect = np.zeros(frames - 1)
    for k in range(frames - 1):
        for i in range(div):
            for j in range (div):
                x = video[k][i*xbloco:(i+1)*xbloco, j*ybloco:(j+1)*ybloco]
                y = video[k+1][i*xbloco:(i+1)*xbloco, j*ybloco:(j+1)*ybloco]
                B = np.sum((x-y)**2)
                if(B > T):
                    count = count + 1
        vect[k] = count
        count = 0

```

Figura 2: Algoritmo para o cálculo do SSD(Square sum difference).

Tal que, T é o limiar testado e div é o tamanho dos blocos (no caso foram usados 8 e 16).

3.3. Diferença de Histogramas

A diferença entre histogramas, foi calculada a partir de dois frames subsequentes, onde encontrava os histogramas de cada um e após isso, somava-se o módulo da diferença de cada nível de cinza correspondente. O algoritmo que contabilizava esse fator era tal:

```

def hist_dif(video):
    frames = video.shape[0]
    vect = np.zeros(frames-1)
    for k in range(frames-1):
        sum1 = 0
        hist1,bins = np.histogram(video[k].ravel(), 256, [0, 256])
        hist2,bins = np.histogram(video[k+1].ravel(), 256, [0, 256])
        a = abs(hist1 - hist2)
        sum1 = np.sum(a)
        vect[k] = sum1

    return vect

```

Figura 3: Algoritmo para o cálculo da diferença entre dois histogramas

Após contabilizar todas as diferenças para cada frame, um limiar T era calculado. A fórmula geral para o seu cálculo é dada por: $T = \mu + \alpha \sigma$, onde μ é a média e σ o desvio padrão da diferença de histograma de todos os frames. O valor de α em geral, varia entre 3 e 6.

3.4. Diferença de Mapa de Bordas

A diferença entre mapas de bordas, foi contabilizada para dois frames subsequentes e após isso, contou-se o número de pixels de borda que diferiam uns dos outros. O mapa de borda utilizado foi **Canny**. Após a contagem, testou-se os valores com um limiar T , para identificar quais frames eram uma transição abrupta. O algoritmo utilizado foi:

```
def edge_det(video):  
  
    frames = video.shape[0]  
    vect = np.zeros(frames-1)  
    for k in range(frames -1):  
        edge = feature.canny(video[k], sigma = 0)  
        edge1 = feature.canny(video[k+1], sigma = 0)  
        x = np.count_nonzero(edge != edge1)  
        vect[k] = x  
  
    return vect
```

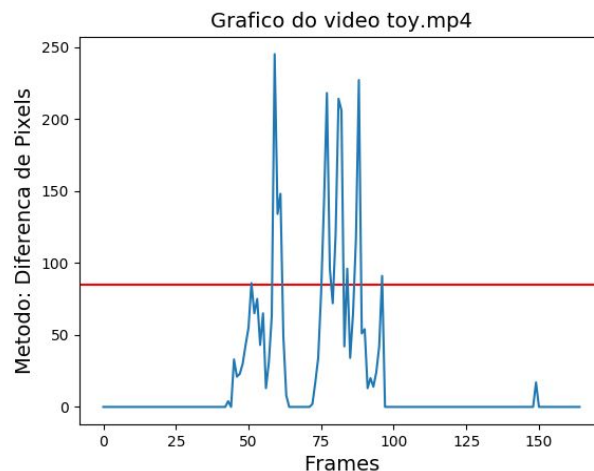
Figura 4: Algoritmo para o cálculo da diferença de borda de um vídeo

4. Análise e testes

Foram utilizados 4 vídeos para teste. Tais eram: *toy.mp4*, *umn.mp4*, *lisa.mpg* e *xylophone.mp4*. Para cada vídeo, gerou gráficos para os 5 métodos citados e após isso, pode-se utilizar limiares para definir quais eram os pixels de transição.

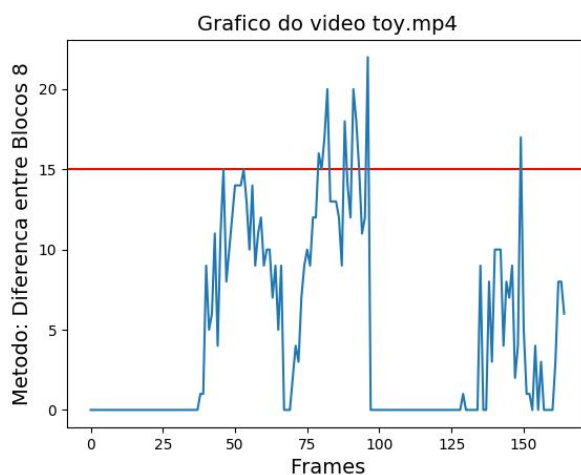
4.1. Toy.mp4

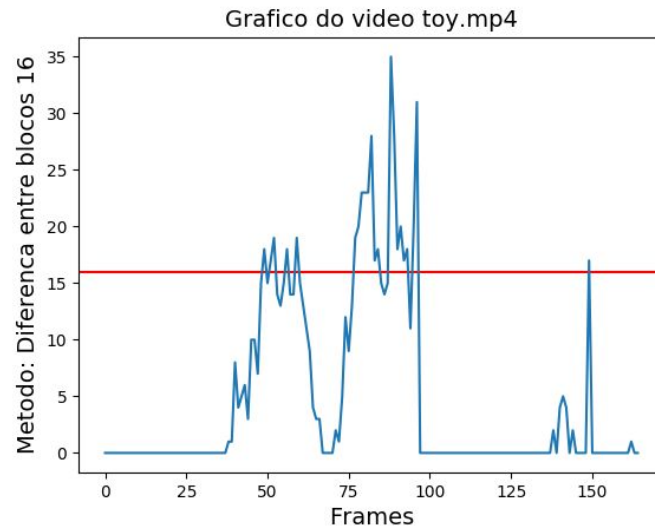
Primeiramente foi gerado um gráfico para a diferença de pixels. O limiar T_1 utilizado foi de 115 para a diferença de pixels. Obteve-se:



A partir dele, pode-se observar que os picos foram os pixels de transição. Como por exemplo, o pixel 76, no qual há uma grande mudança em seu valor em relação ao 75. Portanto, definiu-se o limiar para este sendo: $T_2 = 85$. A linha vermelha indica cada limiar escolhido.

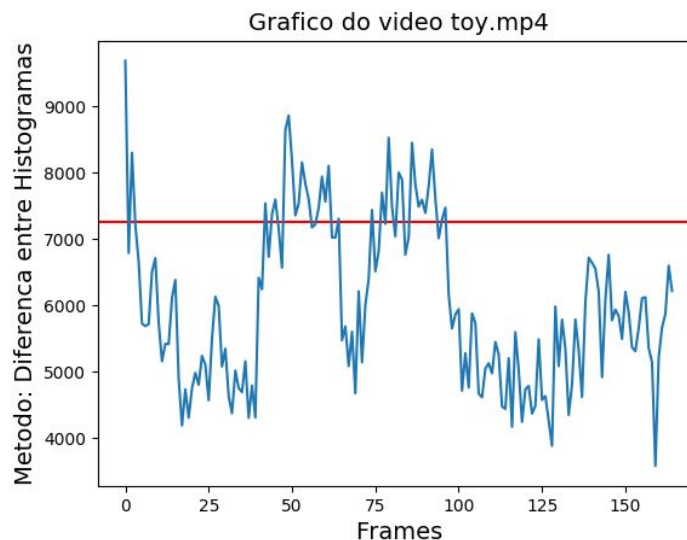
Após isso, foi feita a diferença entre blocos. Primeiramente para 8x8 blocos e depois para 16x16 blocos. O limiar T_1 utilizado foi de 150.000 para ambos, obtendo o gráfico:



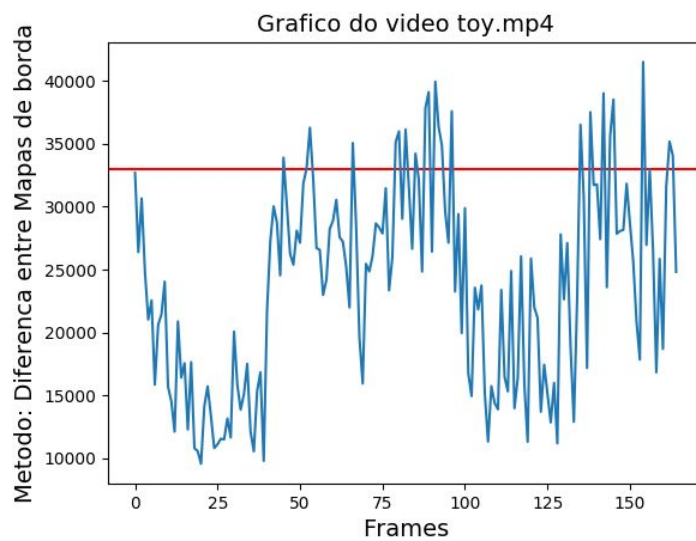


A partir desses gráficos, obteve-se o limiar para 8x8 tal que: $T_{8 \times 8(2)} = 15$. E para o de 16 blocos: $T_{16 \times 16(2)} = 16$, de forma que o picos indicassem transições abruptas. A linha vermelha indica em cada um dos gráficos o limiar escolhido.

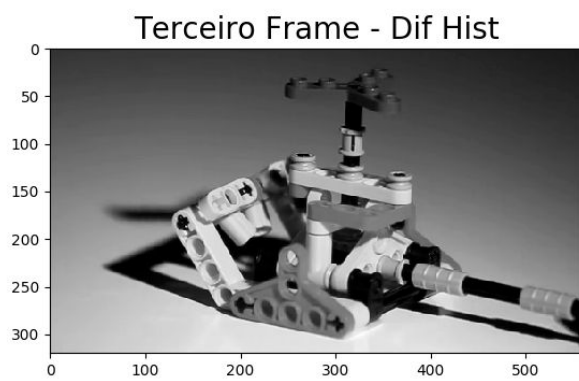
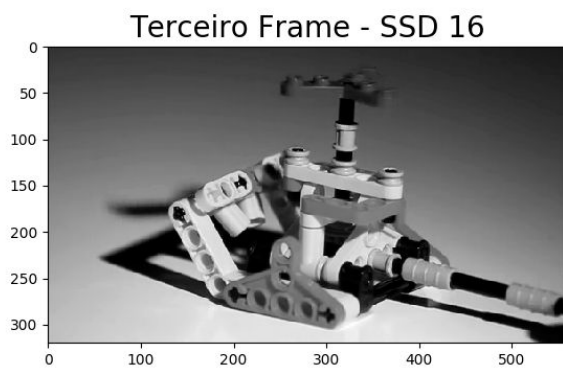
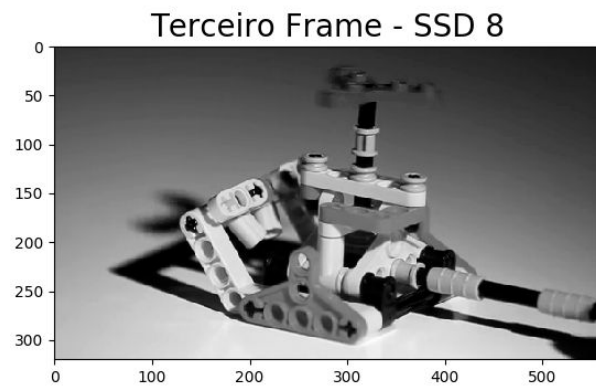
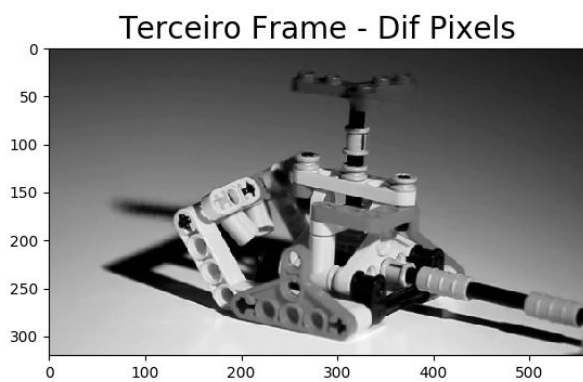
O terceiro método avaliado, foi a diferença entre histogramas. Na qual, obteve-se o limiar utilizando a fórmula proposta na seção [3.3]. O valor de alpha, geralmente varia entre 3-6, porém para esses valores o método não conseguiu obter nenhum pixel de transição, assim o valor foi de 1, obtendo o limiar: $T_1 = 7250$, como observado em vermelho no gráfico abaixo:



E por fim, foi feita a diferença entre mapas de borda. Este, por sua vez, gerou gráficos mais complicados para a observação de transições, porém, foi setado um valor de limiar sendo: $T_1 = 33000$, obtendo o gráfico (em vermelho o limiar):



E assim, obtendo alguns frames:



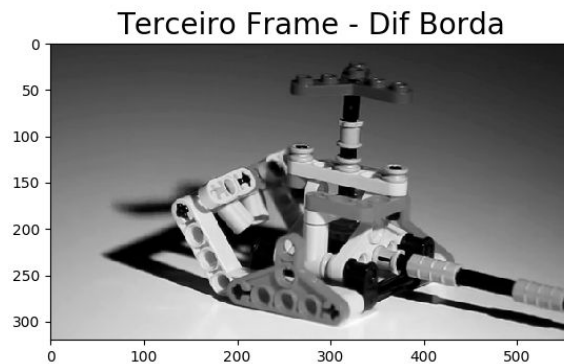


Figura 5: Terceiros frames de transição

Todos os frames acima, foram os terceiros frames encontrados, ou seja, foram o terceiro frame que indica uma transição abrupta para cada um dos métodos. Podemos observar, que os métodos não retornam os valores exatos, mostrando que houve diferenças entre os métodos ao achar frames de transições diferentes. Isso se dá pelo fato de que quando é utilizado o limiar, como nos gráficos acima, não demonstra um valor de transição sempre, uma vez que não só os frames de transição que são considerados e sim todos acima desse limiar.

Resumo dos limiares definidos para encontrar os frames de transições abruptas.

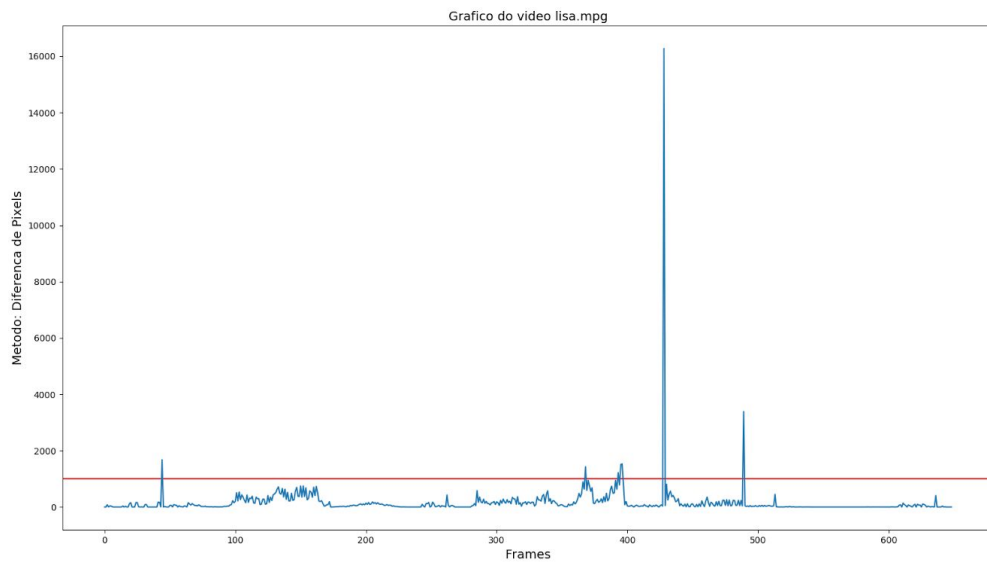
Tabela 1: Limiares utilizados para o vídeo toy.mp4

T2 - Pixels	T2 - SSD 8	T2 SSD 16	T1 - Histograma	T1 - Bordas
85	15	16	7250	330000

4.2. Lisa.mpg

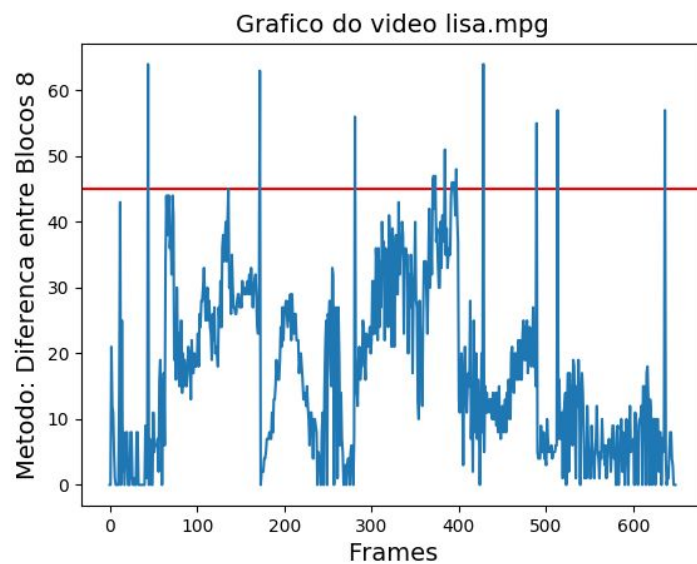
A sequência de execução dos dados para o vídeo "lisa.mpg" foi a mesma para o vídeo anterior.

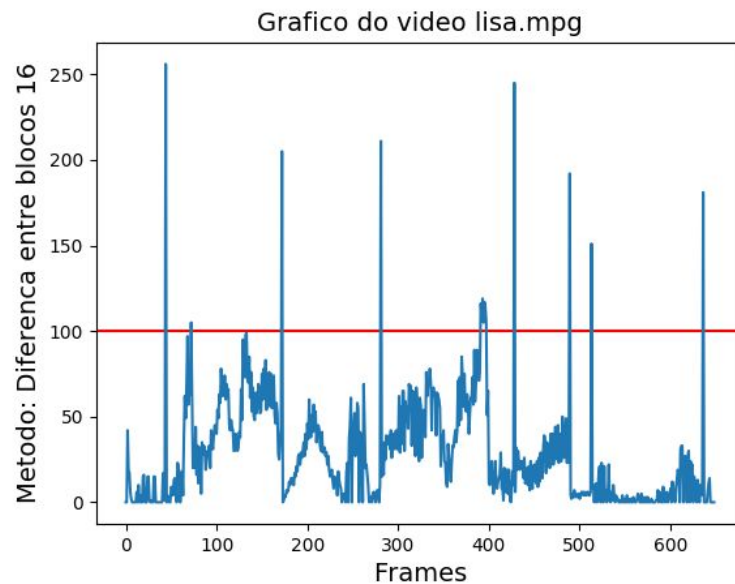
Primeiramente, foi executado a diferença de pixels, como limiar $T_1 = 115$, obtendo o gráfico:



Tal que, o limiar $T_2 = 1000$ (em vermelho) e assim obteve-se os pontos considerados transições abruptas.

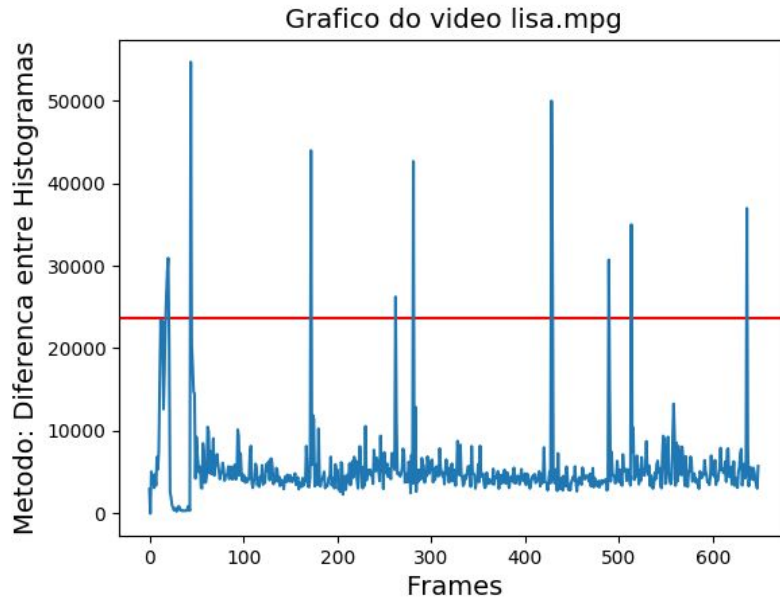
O segundo método, foi a diferença de blocos. O limiar $T_1 = 100000$ foi utilizado para definir blocos diferentes, assim, obteve-se os gráficos para 8x8, 16x16, respectivamente.



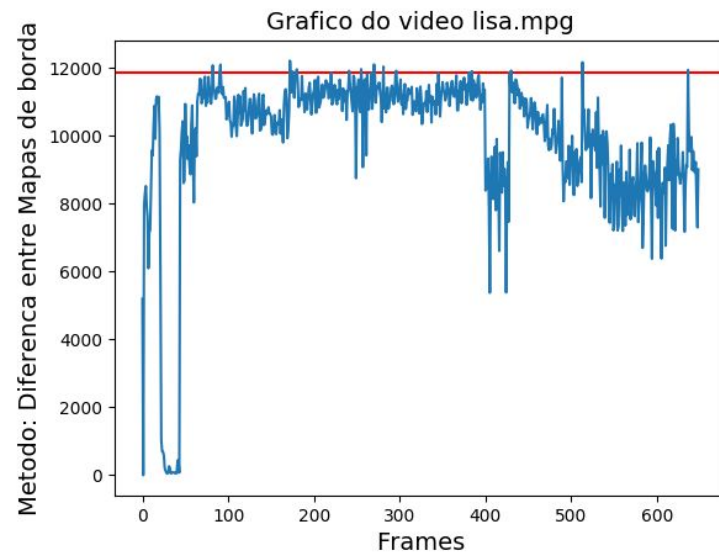


Os limiares foram para 8x8: 45 e para 16x16: 100, podendo observar os picos como transições abruptas dos frames.

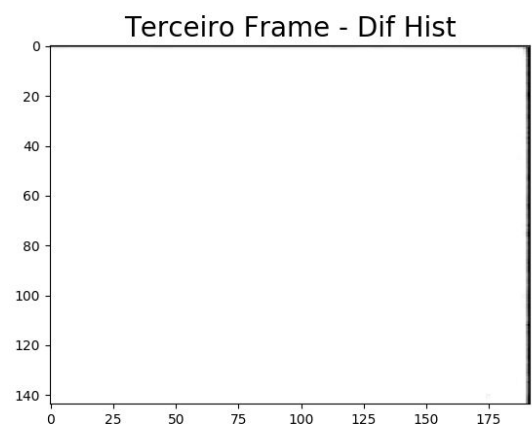
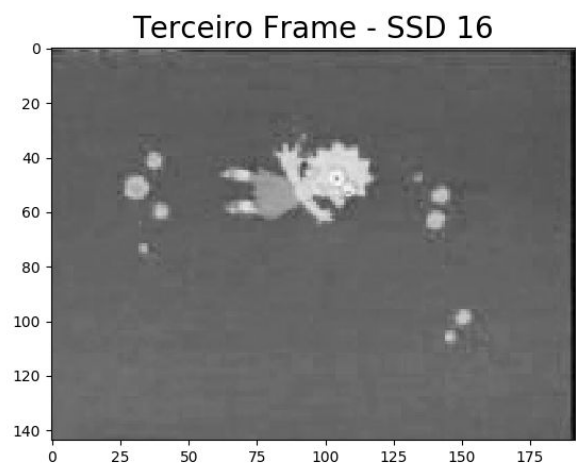
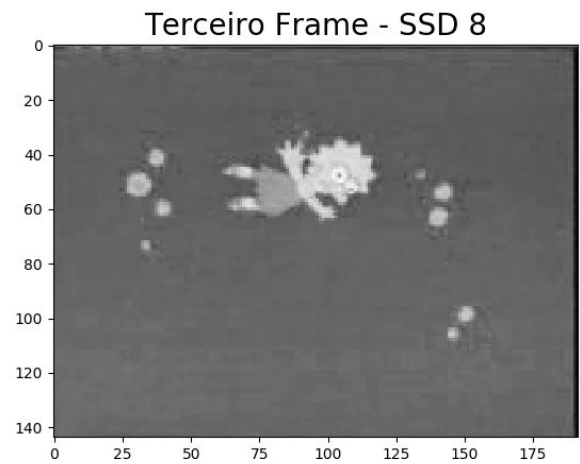
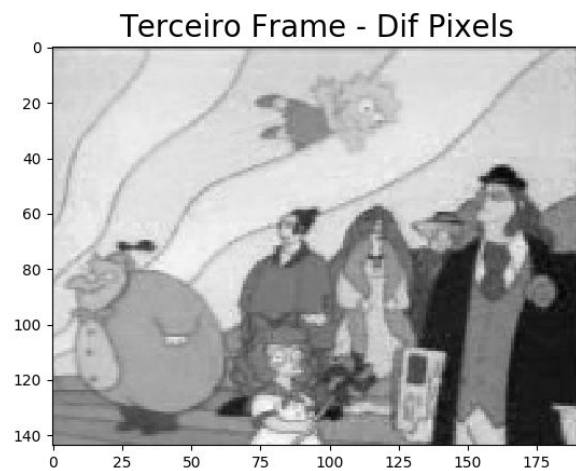
Já para a diferença de histogramas, o valor do limiar foi de 23650, obtido pela fórmula citada na seção 3.3. O valor de alpha foi de 3.5.



E por fim, gerou-se o gráfico da diferença de borda, no qual, também foi um pouco mais confuso de observar as transições pelo forma a qual se desdém. O limiar utilizado foi: 11850



Os terceiros frames encontrados por cada método são:



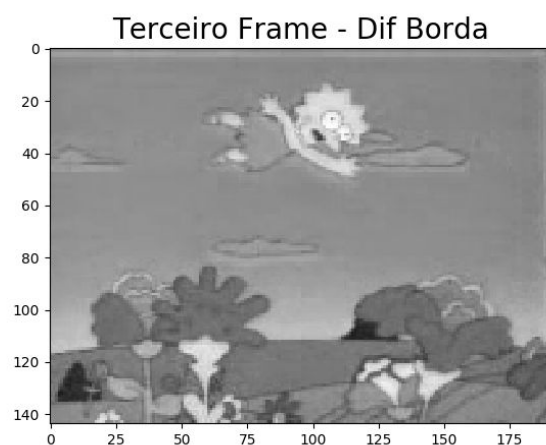


Figura 6: Terceiros frames de transição

Mesmo com os gráficos mostrando os mais definidos em relação ao vídeo *toy.mp4*, os terceiros frames de transição encontrados diferenciam-se em quase todos os casos, salvo, o método de borda para 8 e 16 blocos que são o mesmo. Isso indica que ao traçar o um limiar para achar os picos, não necessariamente estará achando e sim somente os pixels que estão acima. Transições abruptas são dadas por grandes derivadas (grandes variações), ou seja, seria necessário um métodos mais eficientes para encontrar os valores esperados.

Resumo dos limiares definidos para encontrar os frames de transições abruptas.

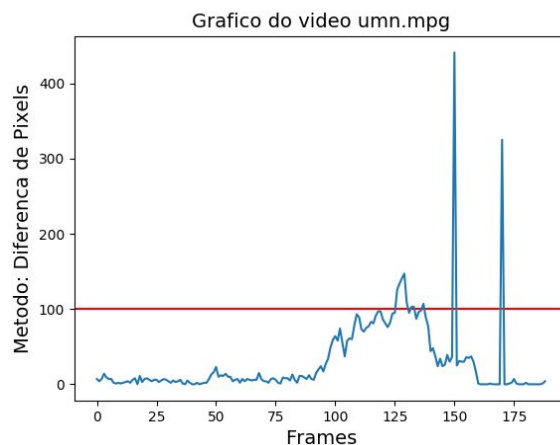
Tabela 1: Limiares utilizados para o vídeo lisa.mp4

T2 - Pixels	T2 - SSD 8	T2 SSD 16	T1 - Histograma	T1 - Bordas
1000	45	100	23650	11850

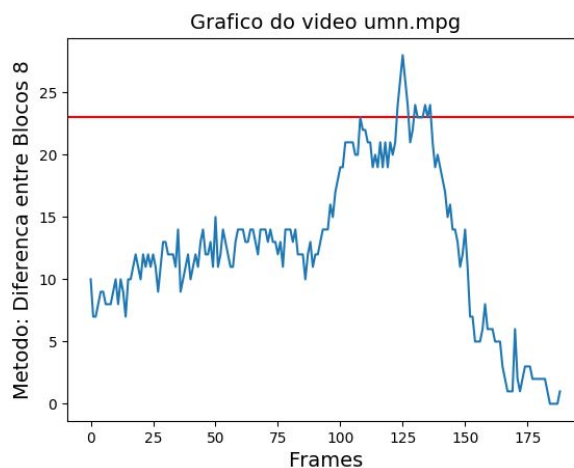
4.3 Umn.mp4

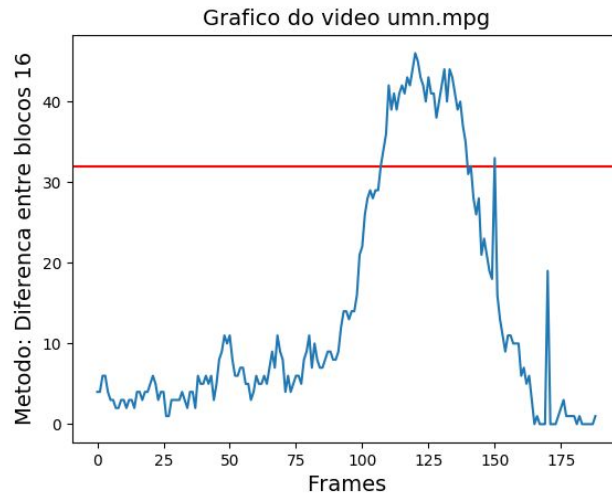
Para o vídeo *umn.mp4*, foram executados os mesmos métodos anteriores, tais como:

Diferença de pixels, foi utilizada tal que $T_1 = 115$ e $T_2 = 100$ (em vermelho), obtendo o gráfico:

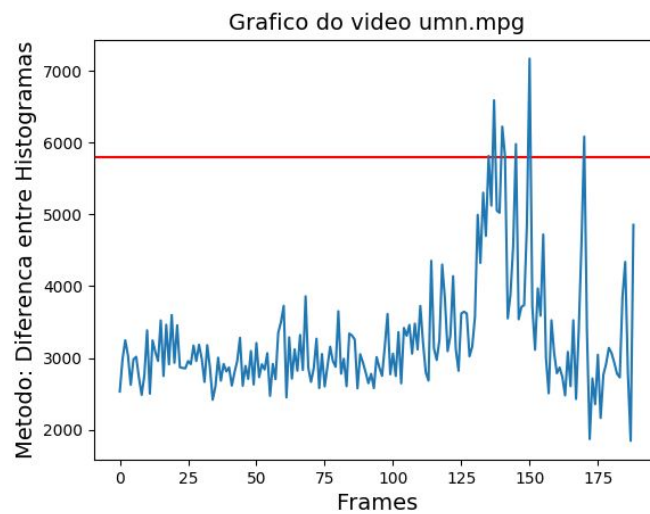


Logo em seguida, obteve-se os dados para as diferenças de blocos. O primeiro limiar $T_1 = 10000$, foi utilizado para encontrar os blocos mais diferentes, depois foi utilizado um limiar $T_{2(8 \times 8)} = 23$ e $T_{2(16 \times 16)} = 32$ para encontrar os frames com transições abruptas. Assim:

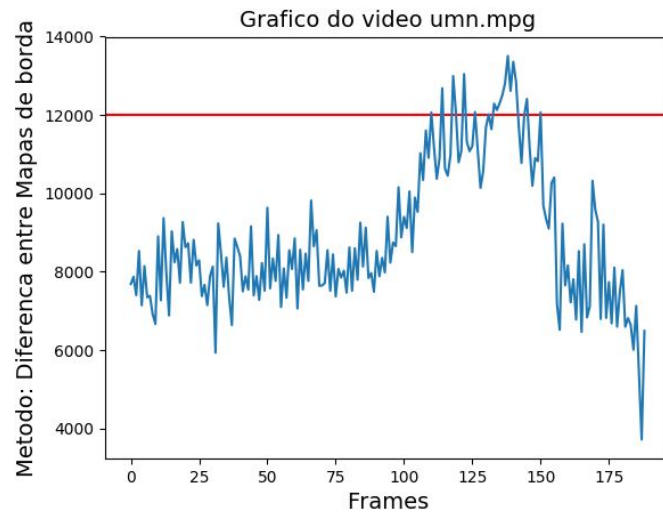




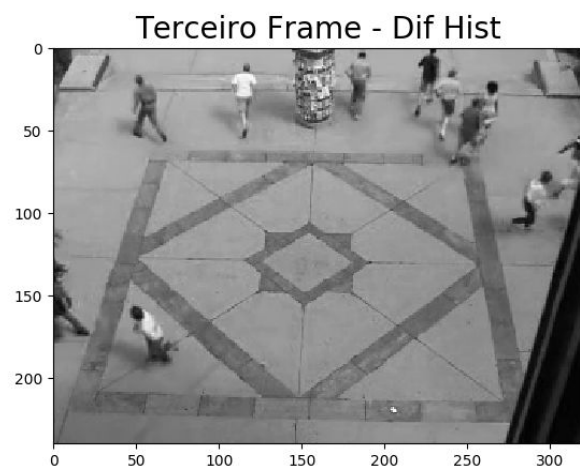
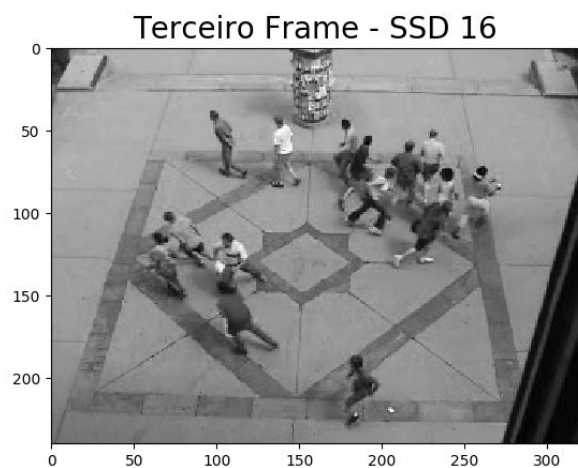
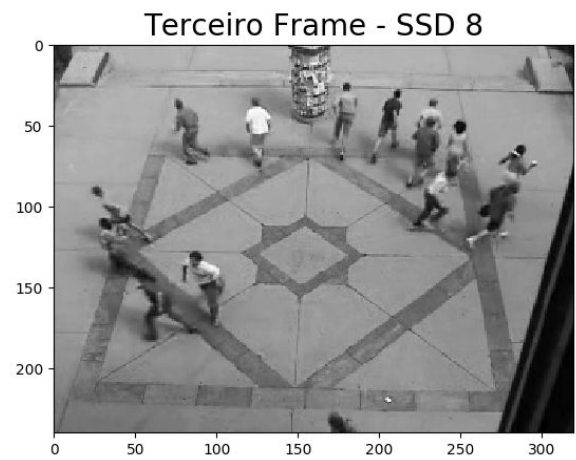
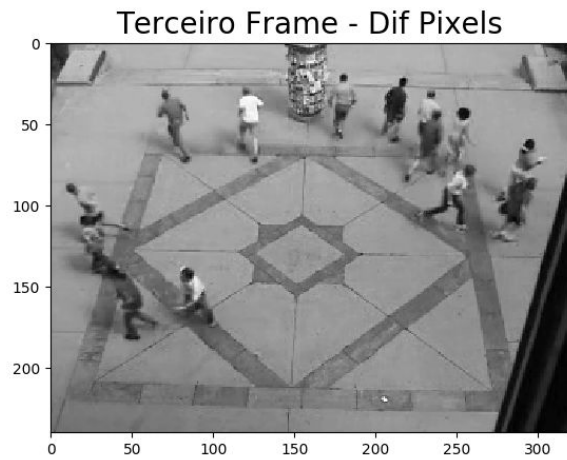
Para a diferença de histogramas, o limiar encontrado com o valor de alpha sendo 3 foi: $T_1 = 5800$, e portanto o gráfico:



E por último, para a diferença de bordas, definiu-se o limiar $T_1 = 12000$ para encontrar as transições abruptas, tal que:



Obeve-se assim os terceiros frames de cada método:



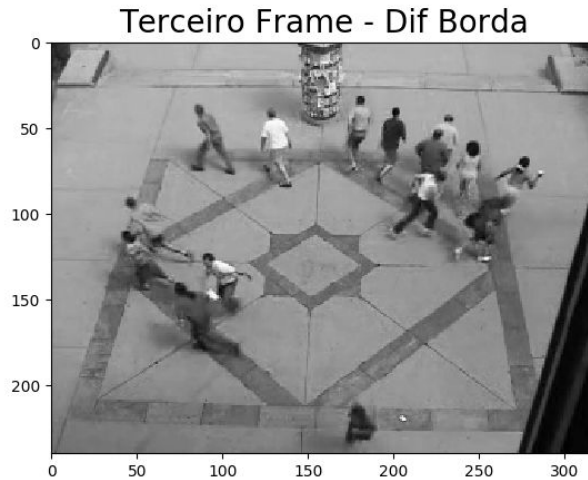


Figura 7: Terceiros frames de transição

Para o vídeo *umn.mp4* observa-se que os terceiros frames de transição para cada método foram mais próximos, uma vez que a imagem possui menos transições abruptas, permitindo com que os algoritmos encontrassem às transições somente nas regiões. Porém, mesmo assim, não são iguais, mostrando a diferença na definição dos limiares e de cada método.

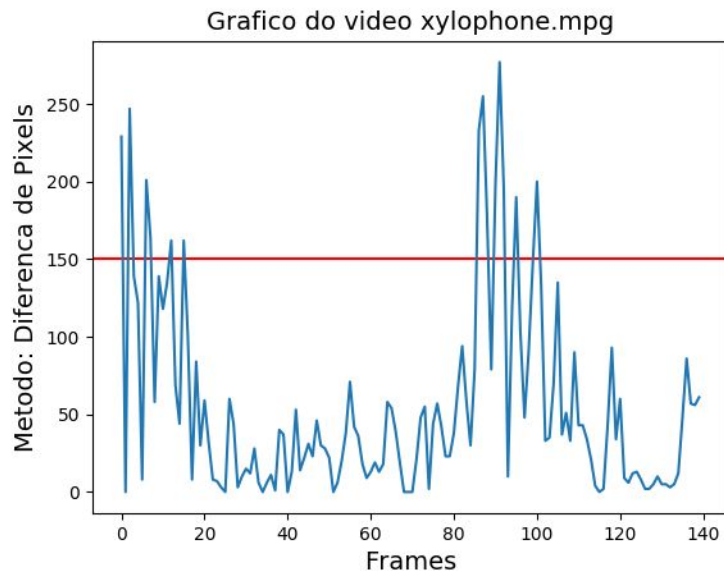
Resumo dos limiares definidos para encontrar os frames de transições abruptas.

Tabela 1: Limiares utilizados para o vídeo umn.mp4

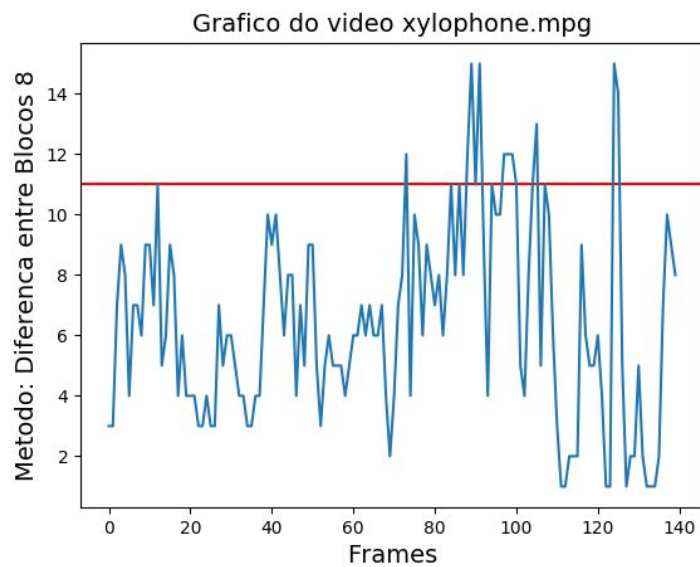
T2 - Pixels	T2 - SSD 8	T2 SSD 16	T1 - Histograma	T1 - Bordas
100	23	32	5800	12000

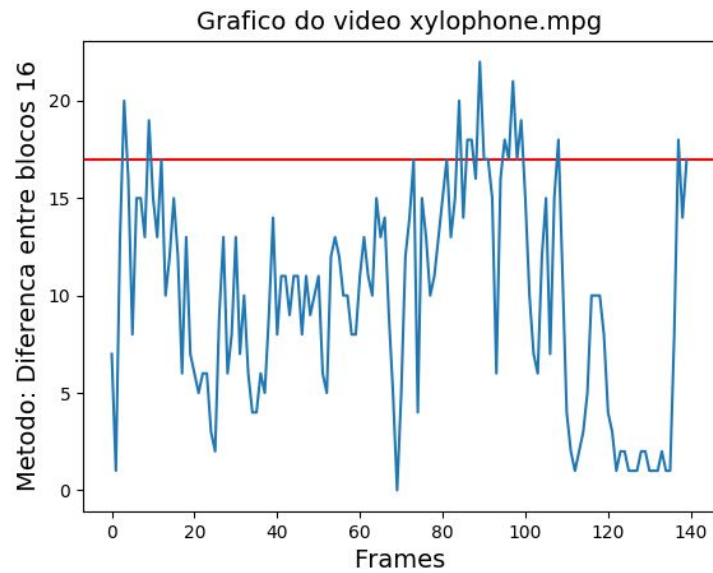
4.4 Xylophone.mp4

Os mesmos procedimentos foram realizados para o vídeo *xylophone.mp4*. Primeiramente, para a diferença de pixels, definiu-se o limiar $T_1 = 115$. Após isso, utilizou-se o gráfico gerado para observar o limiar $T_2 = 150$, obtendo assim o gráfico:

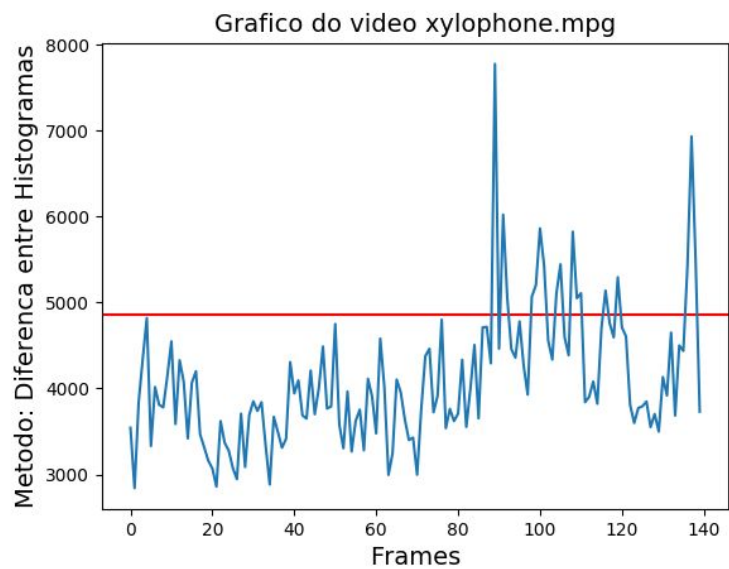


Já para os blocos, o limiar $T_1 = 11000$ foi utilizado para poder encontrar os blocos que mais diferenciam-se uns dos outros. Ao observar os gráficos, obteve-se os limiares: $T_{2(8 \times 8)} = 11$ e $T_{2(16 \times 16)} = 17$. Logo, os gráficos:

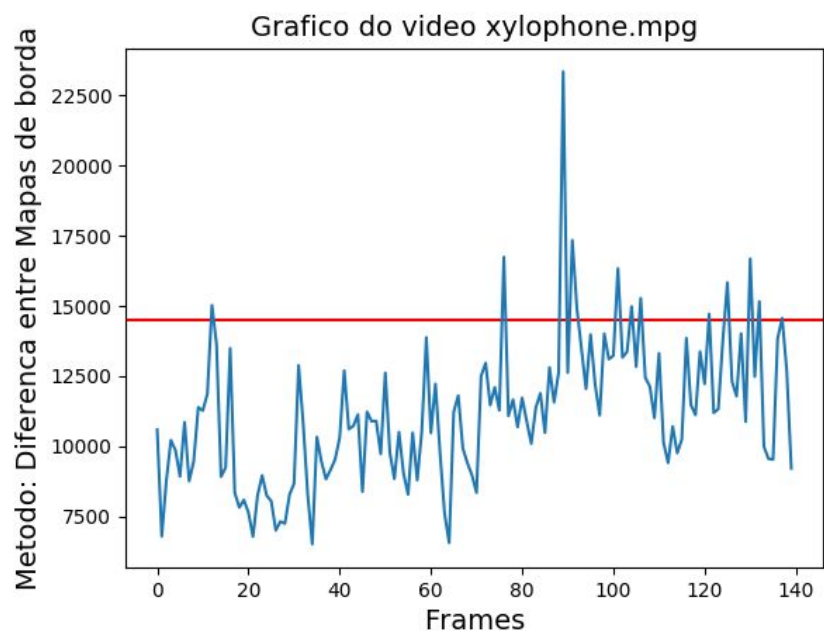




Para a diferença entre os histogramas, calculou-se as somas das diferenças e encontrou o valor do limiar, tal que: $T_1 = 4850$, obtendo assim:

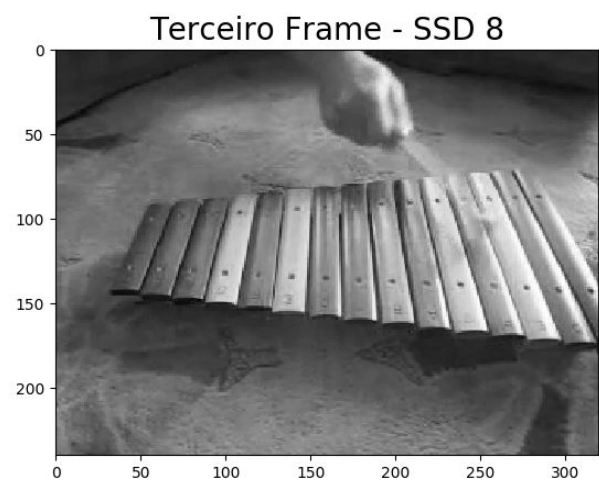
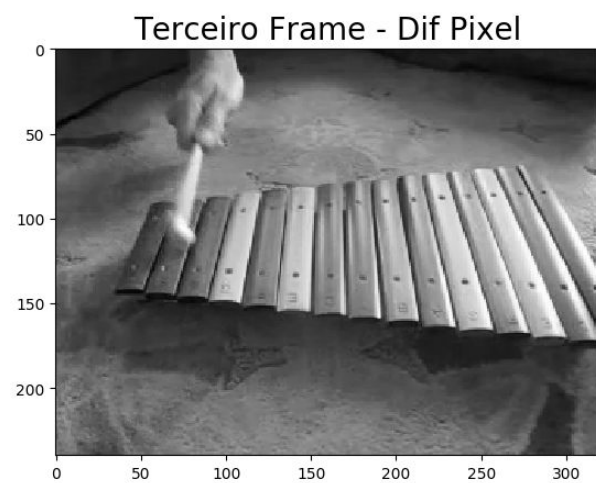


E para os mapas de bordas, utilizou o limiar $T_1 = 14500$, obtendo o gráfico:



Tais limiares, (linha horizontal nos gráficos) foram utilizados para definir transições abruptas.

Os terceiros frames de transição encontrados por cada método são:



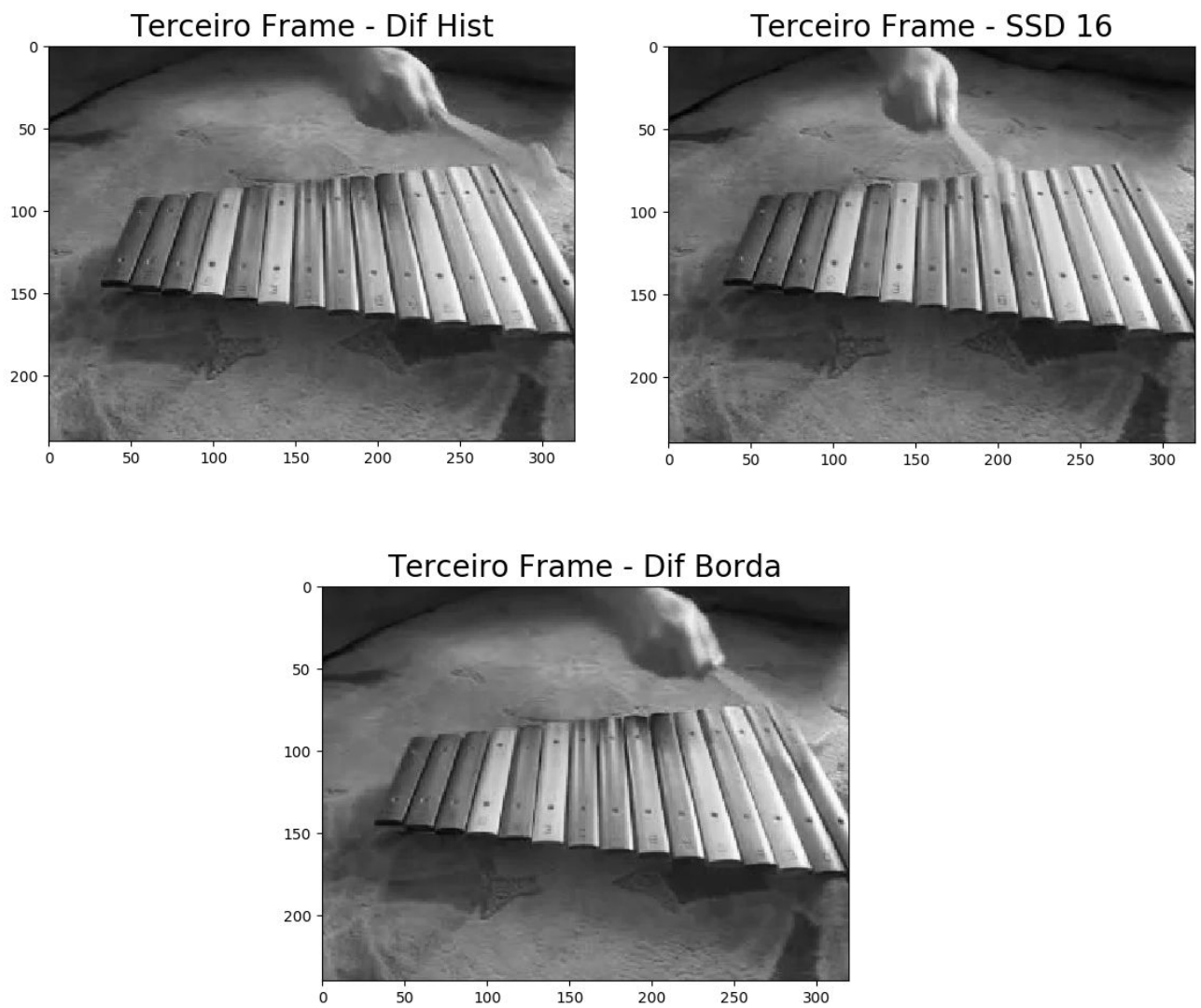


Figura 8: Terceiros frames de transição

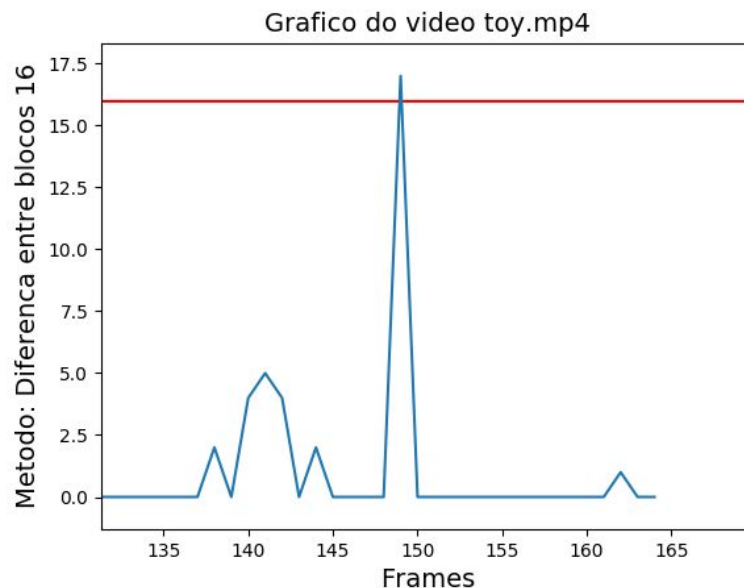
O mesmo ocorre nesse caso, os pixels encontrados diferem uns dos outros pelas diferenças entre métodos e pela dificuldade de definir transições abruptas simplesmente comparando ao limiar.

Resumo dos limiares definidos para encontrar os frames de transições abruptas.

Tabela 1: Limiares utilizados para o vídeo umn.mp4

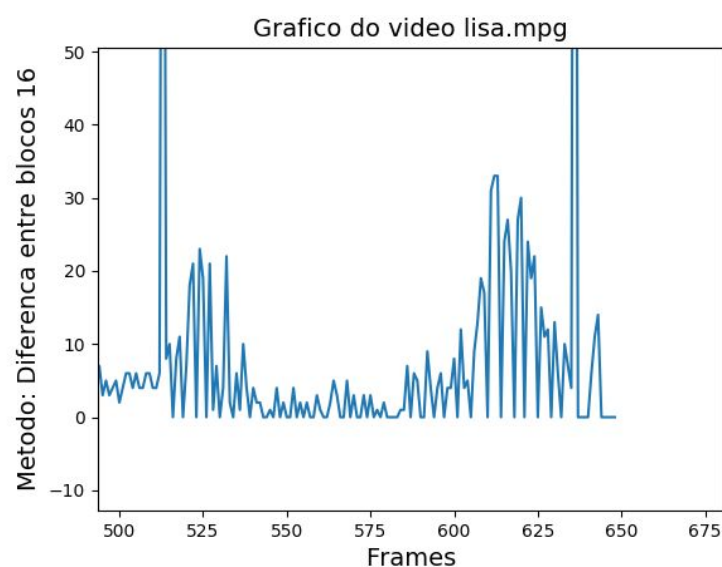
T2 - Pixels	T2 - SSD 8	T2 SSD 16	T1 - Histograma	T1 - Bordas
150	11	17	4900	14500

Observando as diferenças entre cada método em cada vídeo, observou-se que simplesmente a comparação com um limiar não é uma boa definição para encontrar transições abruptas. É necessário um método mais sofisticado no qual encontre os picos (máximos), porém que levam em consideração a relação com os pixels anteriores. Uma transição abrupta define-se por ponto de máxima derivada, em relação aos frames adjacentes, por exemplo, no gráfico de de blocos 16x16 para o vídeo toy.mp4, nos frames finais:



Neste gráfico, pode observar que no pico do frame 149 a sua diferença para os seus adjacentes (148, 150) é grande, ou seja, define uma transição abrupta.

Agora, observando o gráfico abaixo, para blocos de 16x16 do vídeo lisa.mpg nos frames finais:



:

Entre os frames 575 e 630 ocorrem muitas transições, porém, não pode-se afirmar como transições abruptas, uma vez que a diferença para os pixels adjacentes não é tamanha a definir-se uma transição abrupta, uma vez que essa esteja "acima" do limiar.

5. Conclusão

O resultado atingido foi o esperado, mesmo com os problemas na definição para encontrar transições abruptas utilizando somente limiares, chegou-se ao resultado que encontra frames de transição.

Assim, mostra-se que para a obtenção desses frames é necessário, em sua maioria, métodos mais eficientes ou a combinação desses métodos para que possa obter um resultado mais consistente e coerente.