Video Forensics: Can Motion History Images be used to detect doctored videos?

F. Lotfi

1. INTRODUCTION

I debated calling this paper: "Did Acosta accost that intern?" But despite how funny that sounds to me its not the point of this paper. We have a growing trend of doctored and fake images bombarding us but not many tools to detect the fakes. The era of deep fakes that's coming will only make this problem worse. As a society we need tools to drudge through all the misinformation. I will try to answer a single question, was the video posted by the press secretary a fabrication? While this paper covers a recent political event I will not jump into any partisan questions like was the press secretary aware of the fabrication or if Acosta had attempted to hurt the intern. For both questions I have no answers for you, so if you are hoping to find that hear please look elsewhere.

2. OBJECTIVE

- To capture the motion history image (MHI) of the Info-Wars video posted by Sanders.

- To capture the MHI of the original un-edited CSPAN video.

- To overlay the two MHI values captured on two different color channels.

3. HYPOTHESIS

My hypothesis is that the if the videos are the same the MHI values in the blue and red color channels will overlay on one another and we will get a purple MHI image. However if the info wars video is sped up we will see a long red-shift.

4. BACKGROUND

Motion History images were used because they are great ways to get classifications about motion. MHI will tell you where the motion is happening, its intensity (speed), and it's direction (the path). It is used by machine learning and computer vision researchers trying to classify motions such as waving some one towards you vs shewing them away. It can do this bit a temporal algorithm that degrades intensity overtime. In other words, more recent actions have higher intensity and faster actions cover more distance of the matrix and spread that intensity farther.
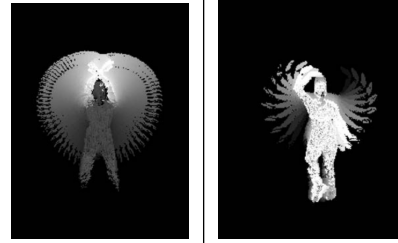

Figure 1: Example Motion History images

5. MHI IN PRACTICE

Essentially we want to diff the two videos, but we can't things like video quality, general noise, banners, and ticket streams are going to make that difficult. So instead the plan is to diff the Motion.
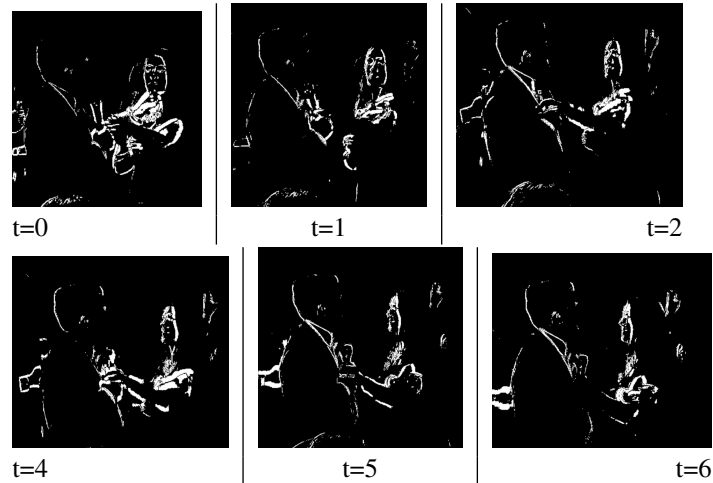

Figure 2: Frames of an MHI stream.

5.1 Does Diffing tell much?

As a visualization showing the diff can help our human eyes tell if the similarity of two videos. If we wanted to go further we could use MHI images as global space-time âĂĲshapeâĂİ descriptor From there we could build a classifier and use a ML algorithm like K-nearest neighbor or support vector machines to compute the similarity. The baseline would be the un-doctored video in that case and I would expect something close to exact match.

5.2 Was MHI the best approach?

Honestly, I'm not sure. I've used them before to write classifiers for motion, never as a technique to detect fraudulent behavior, but this seems like something it could maybe do.

5.3 Results

First did purple correspond to matching frames? I have to say it did, but I'm just eye-balling it.
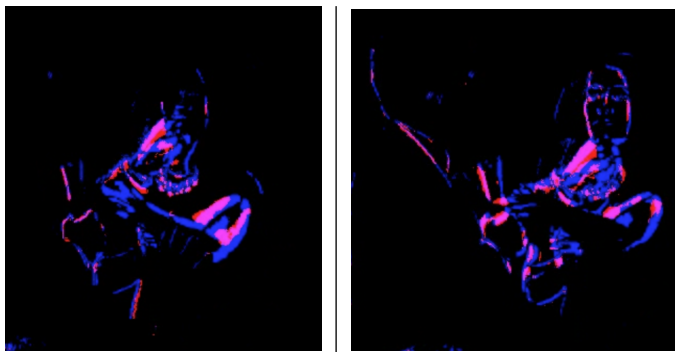

Figure 3: Example of Matching Frames

Did you see anything weird? Yes around the 1 second mark the motion in the Info Wars video just disappears. I've check this isn't a programming bug its the same frame as the previous frame so there is no motion. It is likely no mistake that this motionless frame is picked right before the point of impact. I can only surmise that this may have been a way to make the drop of the hand seem more dramatic. However, that's probably also why so many others called this video out as a doctored by just eyeing the video as it makes the video seem choppy. A better fake probably would have tried to be more fluid.
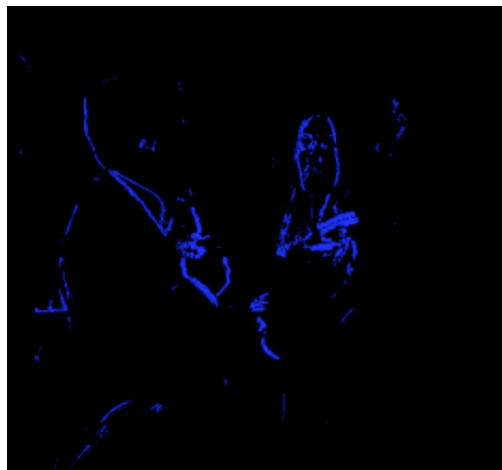

Figure 4: The MHI for the Info Wars video disappears on this frame.

Finally did I see Any red shifts indicating more speed? Yes I did however not How I expected it. I expected it to be a long continuous red shift however it was patchy.
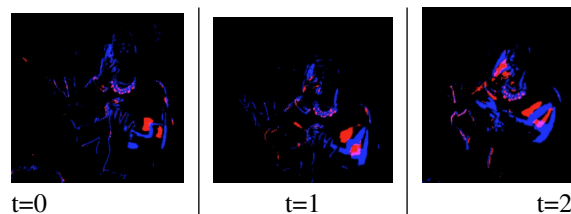

| t=0 | t=1 | t=2 |

Figure 5: concatenated motion history between two video stream.

Looking at the raw video frames shows frames that look blended. I think it has an effect to make the video appear to have more motion than existed, but I don't know if this was an intentional effect or not. As a result the MHI looks like two parallel arms are reaching for Acosta's microphone. Atm not sure if this is an artifact of twitter doing weird compression on its video and so their are not as many I-frames or if there is something else going on here.
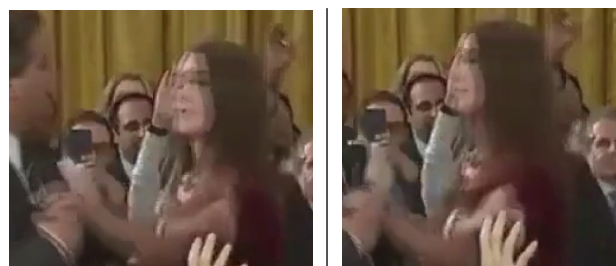

Figure 6: raw frames that appear blended.

6. FUTURE WORK

I don't plan on working on this any further, but I do think you can use something like Hu moments to build good descriptors for a classifier on motion. Maybe if we could automate the detection of doctored videos with ML the public would have a defense against all the miss information bombarding us.

7. CONCLUSION

In both video sources the motions come out great and are prominent. Even with the blended frames from the Info Wars video the motions are mostly unaffected. You can tell direction and intensity easily. As for my hypothesis I think parts turned out to be true others partly so. I think there are strong indicators that this video has been altered, however I think the addition of still frames before impact may have done more to make the drop of Acosta's arm seem more dramatic. I was not expecting that. Further, I did not see the continuous red shift I had expected. Finally, It does appear like some purple frames did exist, which I believe indicates matching frames, this is of note bit is an indicator that these videos are synced up.

8. FEEDBACK

I'd love to consult with expects in this area. Is this a good approach at detecting alterations in video speed?

9. READING MATERIAL

- https://en.wikipedia.org/wiki/Inter_frame
- https://www.pyimagesearch.com/2014/10/27/opencv-shape-descriptor-hu-moments-example/
- https://en.wikipedia.org/wiki/Motion_History_Images