Automated Movie Restoration

Interim Report for Research Project Module 5E1

Xin Shu
Trinity College Dublin
xins@tcd.ie

October 2021

This report is submitted in part fulfilment for the assessment required in 5E1 Research Project. I have read and I understand the plagiarism provisions in the General Regulations of the University Calendar for the current year. These are found in Parts II and III at http://www.tcd.ie/calendar.

1 INTRODUCTION

As high-quality digital movies become more and more prevalent along with the 5-th Generation Mobile Communication Technology, the restoration of archived classic movies helps to address the increasing demand for high quality content. In April 2012, one of the most classic movies, Titanic, was restored into the 3D version and was released on the movie market. This release has obtained more than 350 million USD, according to the box office published by IMDb [1]. Movie restoration not only reproduces the great work with a much higher quality but can also be utilized to rectify the defect frames of today's motion picture industry.

Traditional motion pictures were normally stored on celluloid films. Physical actions on the medium, such as transportation and showing on the motion-picture machine, can potentially cause inevitable and irreversible damage to the pictures. Flicker, "dirt and sparkle", and line scratches are the most common visual artefacts in classic movies [3]. Dirt defects and line scratches are caused by material stuck to the original film and will consequently appear on the digital version after scanning. Therefore, it is important to develop a robust algorithm that is able to remove the artefacts from the original pictures, to reproduce the greatest artworks with higher quality for future generations. This project will focus on solving issues of line scratches. The scratches are normally thin, vertical, semi-transparent lines, appearing across continuous frames at roughly the same location. Figure 1 shows several frames from ***Cinecitta*** shoot in the traditional method. The major research orientation is to detect such scratches and repair the frame whereby the audience can have higher satisfaction.

[1] IMDb pro box office, last update on October 29, 2021.

Figure 1: *Sample taken from frame 35 of the* ***Cinecitta****. The second image is the zoomed frame 35 over its scratched area.*

2 PROBLEM OBJECTIVE

This project will continue the work of previous MAI researcher Warren Pretorius [4], who has successfully built a Deep Neural Network to repair the line scratches of frame. To further develop the algorithm of line scratch detection and removal, the project is divided into several objectives:

- Artificially create degraded video data for training line scratch detection neural network.
- Understand and reproduce the previous algorithm by Pretorius. This will importantly save the time consumption of exploring existing materials and testing.
- Further develop the algorithm so that it is able to handle degraded videos.
- Explore through more network structures and video processing algorithms to enhance the detection performances.
- Conduct subjective tests to evaluate the results.

Extension of the objectives may be carried out to further explore this area of interest if time is affluent.

3 PREVIOUS WORK

Film restoration has been an area of interest to researchers and film directors for more than thirty years. One of the most acknowledged achievements was summarized in Motion Picture Restoration [3], published by Kokaram in 1998. One of the theses from this book profiled the line scratches with a damped sinusoidal

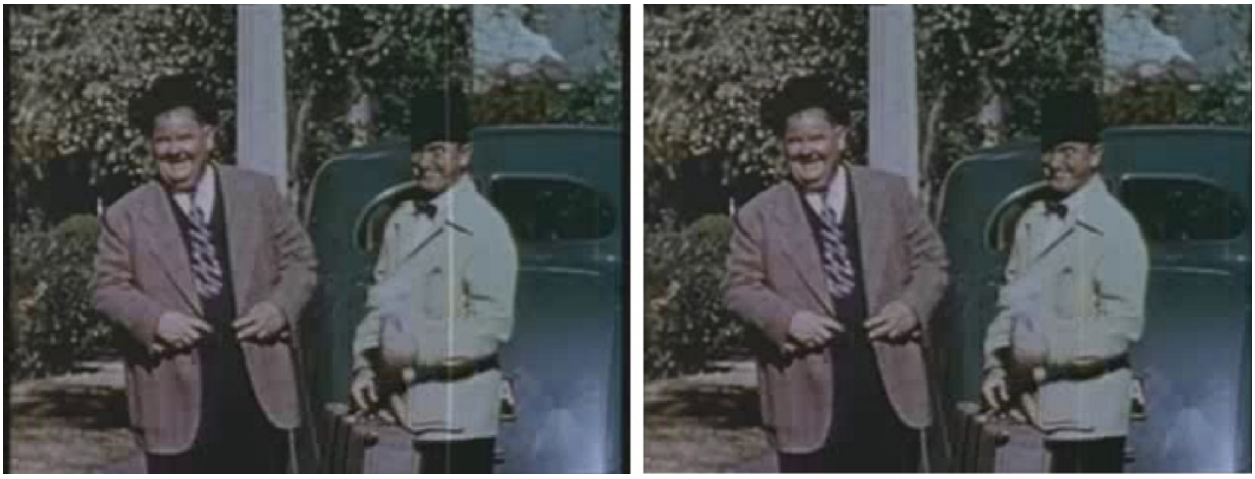

Figure 2: *Restoration results of frame 1 of Laurel & Hardy clip, conducted by Pretorius*

equation:

$$L_n^{(p)}(i,j) = b_p k_p^{|i-(m_p j+c_p)|} \cos \frac{3\pi|i-(m_p j+c_p)|}{2w_p} \tag{1}$$

where:

- $L_n^{(p)}$ is the profile of the pth scratch line.
- b_p is the brightness of the central portion of the line.
- k_p is the decay of the line profile, $k_p \epsilon$ [0.05, 0.95].
- i, j are the horizontal and vertical abscissa.
- m_p, c_p are the slop and intercept with the horizontal edge of the image which defines the orientation of the pth scratch line.
- w_p is the width of the pth scratch line.

Equation 1 helps to artificially create degraded video with the ground truth binary mask. Therefore, the output dataset can be utilized to train a neural network for scratch line detection. Previous work by Pretorius [4] followed this process and achieved decent results of artificial degraded movie restoration. The detection network used Residual Network architecture (ResNet18)[1] as the encoder and its mirror forming the decoder. Once a scratch is detected, the program will empty the pixels on the vertical line and the rest of the image is treated as a missing data problem. This algorithm successfully removed line scratches from the artificial degraded frames.

However, with this method, the restoration program did not perform well on real degraded clips from actual archived material. A frame comparison example is shown in Figure 2 The reason summarized by Pretorius is that either the network was limited by the training set, or the network was not powerful enough to present the best precision. Also, the previous work aimed on restoring degraded frames or short clips, which cannot satisfy the high quality demands of the modern movie industry. With these conditions bearing in mind, the future work can avoid wasting time on the imperfect solution.

In addition, Pretorius also offered invaluable suggestions for carrying out future work and indicated potential methods that can help overcome the drawbacks of neural network. These suggestions will be applied to this project and the results will be included in the final report.

4 ETHICAL

Since the work of this project does not involve medical or health issues, the ethical concerns will hardly be raises during the process. However, sourcing movies and degraded clips will potentially cause copyright concerns. The work content to date has involved using several movie clips. These clips were downloaded from Xiph.Org Foundation, which is a non-profit corporation dedicated to supporting and developing free, open protocols and software to serve the public, developer and business markets. The future work will also be carried out using open-sourced data or obtained from other organizations with valid permission if necessary.

The plan of the work is to artificially degrade a fine video and restore the degraded video back to a similar quality as the original. From the work by Pretorius, the differences between the original and the restored frames were extremely difficult to detect by human vision. To avoid issues of restored vision confounding the original version on the Internet, artificially degraded clip and its restoration should not be published as entire video, but should only be taken several frames to indicate the result of the algorithm.On the other hand, in order to differentiate the real degraded movie and its restoration, the output video of the successful program should be added with a distinct watermark or named as 'restored version'.

In the evaluation stage of the project, subjective tests may require participants to express the visibility of the scratches. The expressions will be utilizes to evaluating the performances of the final algorithm. The entire testing process will be conducted following the GDPR[2]. This project guarantees that the testing process should not obtain any form of personal data of participant at any stage.

[2] EU General Data Protection Regulation.

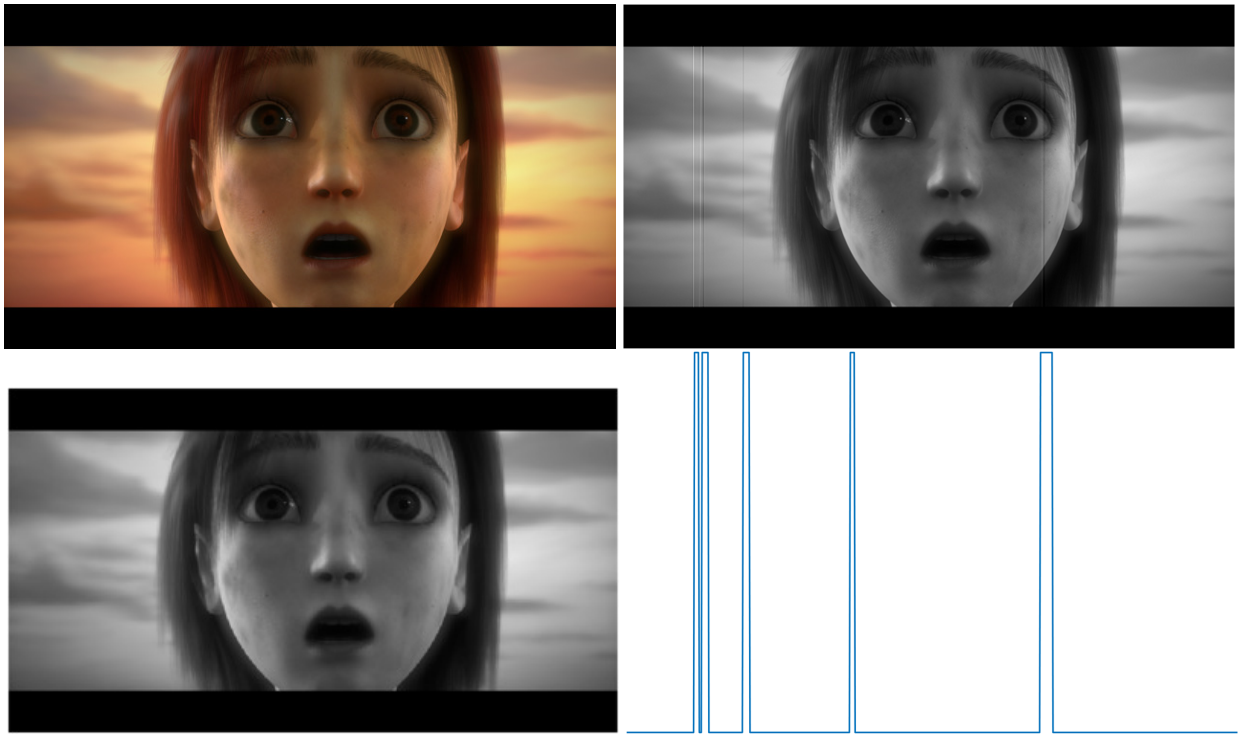

***Figure 3: Top left**: the original frame 671 taken from **Sintel** clip. **Top right**: the artificial degraded frame 671. **Bottom left**: the artificial degraded frame 671 with multiple line scratches. **Top right**: the correspondent binary mask that indicates the horizontal abscissa of the scratch.*

5 PROGRESS TO DATE

Currently, the first objective described in Section 2 is completed. A video clip was obtained from xiph.Org. The clip is called **Sintel Trailer**. The original clip is 53 seconds in length, which contains 1253 frames with 3 RGB channels and the resolution is 1280×720. The degraded clip contains the same number of frames but is in gray scale, the resolution was downsized to 640×360. Modern video processing has better performances with Convolution Neural Network (CNN) involved. Since most of current CNN models are pre-trained with smaller input images, such as 314×314 as input image size of typical **caffe model** [2], it is important to decrease the resolution to achieve similar precision. This will be improved to adapt the cases of larger frame input.

To simulate the line scratches on the scene, random numbers of profiles are created using Equation 1 and the profiles are appended on the original frame at random vertical abscissa to simulate the scratch effect. The degraded frames are finally compressed into an entire video for future processing. The binary mask of each frame that indicates the abscissa of appended is also stored as ground truth for future clas-

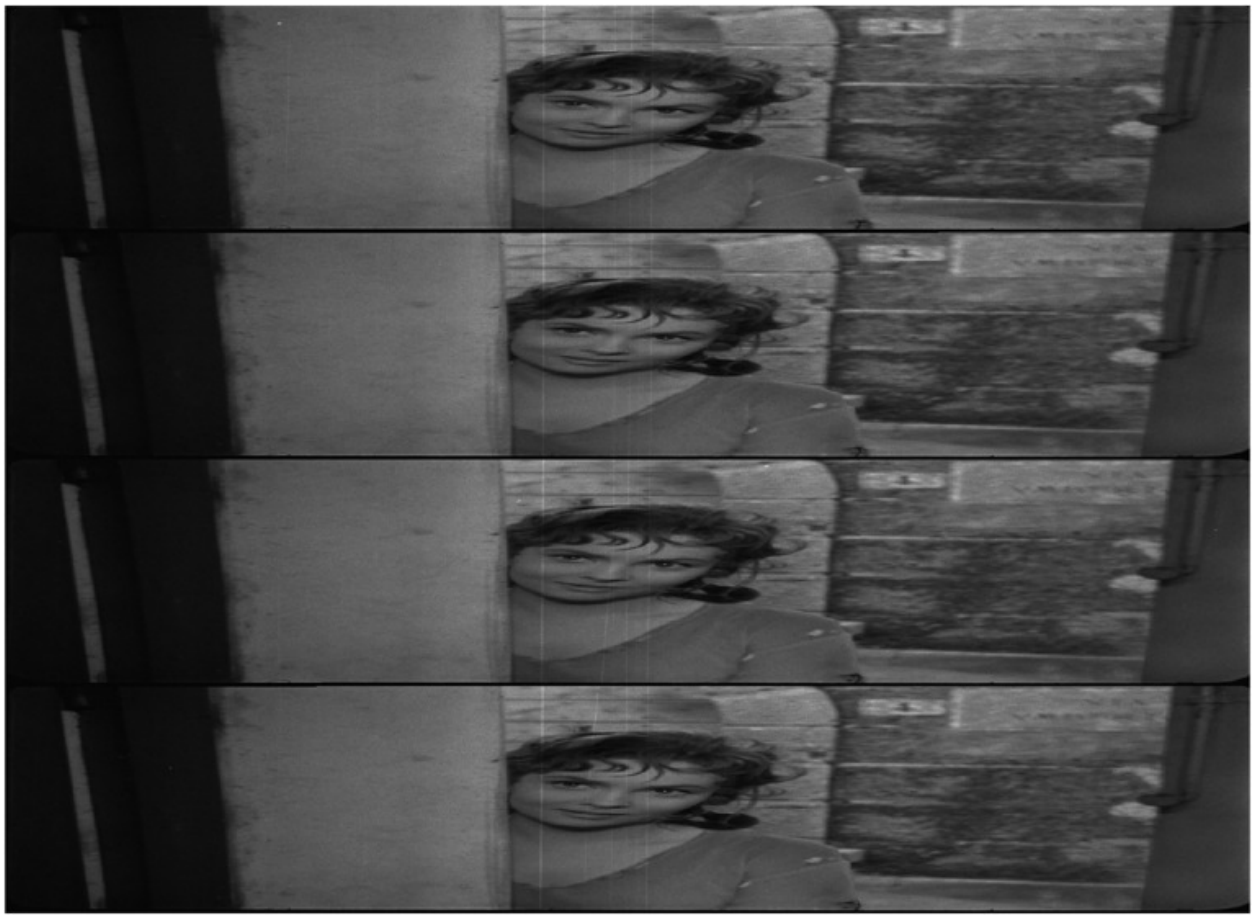

Figure 4: *Compressed frames 21 - 24 taken from* ***Cinecitta****.*

sification training. The example of processing one of the **Sintel Trailer** frames is shown in Figure 3.

For real line scratches in actual degraded films, it is found that the scratches appear in similar columns for several continuous frames. This can be indicated by an example shown in Figure 4. The continuous frames were resized from 1332x1828 to 444x1828 to highlight the scratches and aligned to the bottom of previous frames. In this way, the actual scratches can be more easily classified by networks and are more distinguishable from vertical edges. These methods may also include in the creation of artificial line scratches to help train a more realistic network.

6 SELF-ASSESSMENT

By the midterm of Michaelmas semester, I had made good progress. Adequate materials and literature regarding degraded films and CNN classifiers have been reviewed to grasp a higher-level frame of the solutions. The next stage is to develop the core algorithm for the line scratch detection classifier. If this step can be completed and evaluated by the end of December, more sufficient time is allowed to complete the rest of work as well as further improve the precision. In the meantime, this project provides an excellent opportunity for me to apply the experiences and skills I gained from college lectures, reviewed literature and previous research internship with respect to Machine Learning. Therefore, I am confident that this project can be completed on time.

7 PROJECT PLAN

The subsequent plan of the project is made to follow the timeline of the current academic year structure, started on September 13, 2021. The detailed plan is shown below:

- November and December: Re-implement Pretorius line scratch detector algorithm and evaluate the performance, attempt to discover a possible improvement.
- January: Design an architecture to incorporate motion or multiple frames, explore the feasibility of applying this algorithm into scratch detection. This process requires fair amount of test and potentially iterating over the architecture. Introduce **RAFT** [5] to use explicit motion compensation as a pre-process OR train the motion estimator at the same time as the restoration process?
- February: Start to evaluate the program with subjective tests, for example, visibility testing that requires participants to rate the level of deficiency of the restored video.
- March: Prepare and conduct the final presentation and assess the work.
- April: Write the final MAI project dissertation.

BIBLIOGRAPHY

[1] He, K., Zhang, X., Ren, S., and Sun, J. Deep residual learning for image recognition. *arXiv preprint arXiv:1512.03385* (December 2015).

[2] Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., and Darrell, T. Caffe: Convolutional architecture for fast feature embedding. *arXiv preprint arXiv:1408.5093* (August 2014).

[3] Kokaram, A. C. *Motion Picture Restoration.* Springer, January 1998.

[4] Pretorius, W. Line scratch detection & removal in digital film via deep neural nets. Tech. rep., Trinity College Dublin, April 2020.

[5] Teed, Z., and Deng, J. Raft: Recurrent all-pairs field transforms for optical flow. *arXiv preprint arXiv:2003.12039* (March 2020).