

Towards Automatic Blotch Detection for Film Restoration by Comparison of Spatio-Temporal Neighbours

Peter Gaughran, Susan Bergin, and Ronan Reilly

Department of Computer Science, National University of Ireland, Maynooth,
Co. Kildare, Ireland

{Peter.Gaughran, Susan.Bergin, Ronan.Reilly}@nuim.ie

Abstract. In this paper, a new method of blotch detection for digitised film sequences is proposed. Due to the aging of film stocks, their poor storage and/or repeated viewing, it is estimated that approximately 50% of all films produced prior to 1950 have either been destroyed or rendered unwatchable [1,2]. To prevent their complete destruction, original film reels must be scanned into digital format; however, any defects such as blotches will be retained. By combining a variation of a linear time, contour tracing technique with a simple temporal nearest neighbour algorithm, a preliminary detection system has been created. Using component labelling of dirt and sparkle the overall performance of the completed system, in terms of time and accuracy, will compare favourably to traditional motion compensated detection methods. This small study (based on 13 film sequences) represents a significant first step towards automatic blotch detection.

Keywords: blotch, detection, film, restoration, machine vision.

1 Introduction

Over time, the mechanical viewing of analogue film reels is abrasive and destructive. Even the storage of film reel archives unavoidably suffers because of chemical breakdown such as vinegar syndrome [2]. Given this eventual destruction, saving the material makes its digital conversion essential. Digital conversion also has the comparative ease of distribution of the content once it has been scanned and converted (Digital TV, Blu-ray/DVD discs, downloads). Unfortunately, while the remastering process ensures no further decay will occur, damage incurred prior to scanning will of course be retained.

Traditionally, when restoring footage, each frame of a motion picture reel must be chemically bathed. An average feature length of, for example, 2 hours (7,200 seconds) at 24 frames a second, results in approximately 173,000 frames that have to be cleaned. Specific areas per frame must also be visually identified and dealt with by hand - cleaned if *dirt* (not removed by the initial step) is present, or 'filled in' if *sparkle* is found. *Dirt* blotches are mainly caused by dust or dirt that has adhered to the film stock over time and result in 'dark spots' on a frame, whereas *sparkle* is

caused by the abrasion of the celluloid, resulting in the silver nitrate showing through as 'bright spots' (Fig. 1 shows some examples). Both are random and single frame defects, that is, no specific instance of dirt or sparkle will appear at the same coordinates in a preceding or following frame. It should be noted that line scratches are different in that they manifest as vertical lines, and over several frames will remain in approximately the same spatial position [3]. As a result of this property, line scratches are suitable for detection by the means proposed here.



Fig. 1. A sample frame from *Duck and Cover* (1951). Some examples of dirt are enclosed in the squares, while sparkle is encircled.

Given the time consuming nature of traditional restoration, it is extremely expensive and labour intensive. It is possible to digitally restore the material using various intelligent and/or signal processing algorithms. These algorithms seek to improve the subjective visual quality of the individual frames. Whether digital or more traditionally analogue, an entire film restoration system is usually composed of scene segmentation, flicker correction, blotch and scratch detection and removal, image stabilization and (perhaps most controversially) noise reduction¹. Of these stages, it is the detection of dirt and sparkle blotches that is the focus of this study.

Industrial software exists (such as AlgoSoft, Amped and DIAMANT) – but the means of detection and success rate are unpublished; however, peer assessment and cinematic critique has not been favourable [4]. Previous academic research includes detection of dirt and sparkle by means of motion estimation and 3D autoregressive

¹ Film grain – small grains of a metallic silver halide derivative - may show when projecting a reel if enough photons hit when recording. Many directors deliberately shoot in such a way as to enhance this effect, yet some restoration algorithms remove it. For more information, see [5] *American Cinematographer*, *Post Focus*, May 2008.

modelling – in particular, the JOMBADI (Joint Model Based Detection and Interpolation) algorithm [6]. The JOMBADI approach attempts to combine blotch detection and repair in a single step; a statistical model of the frame is created and motion vectors randomly adjusted until a predicted (reconstructed) frame is reached (based on either prediction error or maximum number of iterations). This results in either very high computational loads and/or lack of accuracy.

Global Motion Segmentation for blotch detection has also been attempted – using this technique, blotches are detected as ‘areas’ of pixels that do not adhere to *any* parametric global interframe transformation model [7]. Being exhaustive, this also results in a high computational load, and is subject to the accuracies, inaccuracies and possible contradictions of the various transformation models employed.

Czúni *et al.* have implemented DIMORF - a neural network for semi automatic detection coupled with an XML database to minimise false positives (by meta tagging incorrect finds in a single frame, all other such instances can be ignored if found in subsequent frames) [8]. As such, DIMORF aspires more to being a semi-automatic detection and indexing software system.

The detection method proposed here is composed of two stages; firstly a single frame is deconstructed into separate components, secondly, these components are then compared with their immediate (if any) temporal neighbours. In order to begin the detection process a chosen frame is converted to two binary images (one based on the original frame for dirt detection, the other an inverted copy, for detecting sparkle). To achieve this goal with minimal loss of component information, local thresholding is applied to the source frame. This thresholding process is outlined in the following section, with the means of blotch detection in single and multiple frames in described Section 3. Section 4 details the experimental setup and results, and Section 5, future work.

2 Preliminary Isolation

In order to separate potential blotches from other objects within a frame, two processes are applied :- (i) local thresholding, and (ii) contour tracing.

2.1 Local Thresholding

Also known as *dynamic* or *adaptive* thresholding [9], local thresholding is used to isolate (but not identify) individual components of objects within the selected frame. Instead of a global threshold applied to all pixels, this operation changes the threshold dynamically across the frame. Developed for industrial use when uneven lighting conditions prevented the segmentation of a lighter foreground object from its background, local thresholding gives binary image output by analysing each pixel with respect to its local neighbourhood. The mean of this neighbourhood is calculated and if the current pixel is black (*i.e.*, an object pixel) it is thresholded to white if the difference between the calculated mean and the current pixel value is lower than that of a user defined offset.

The value of this offset will depend on a variety of factors, including, but not limited to, film grain, differing illumination between scenes and possibly motion.



Fig. 2. The original frame, left, and the same frame after it has been locally thresholded

The values employed in this study were obtained from frame based training sets that were manually derived by the first author. (This manual process is an extremely time-consuming and difficult one, requiring thorough examination of each sequence frame by frame.) These values could then be used across frame sequences containing live action or animation, with differing illumination but were shot on the same film stock (i.e., the same level of grain). Future work will concentrate on the automatic generation of this offset. Figure 2 gives an example of a frame before and after local thresholding. After the thresholding, a contour tracing algorithm is then applied to label all potential blotches.

2.2 A Linear-Time, Contour Tracing, Component-Labelling Algorithm

Originally designed for document analysis and recognition (DAR), the algorithm employed to label the contours and components within the generated binary images was devised by Chang *et al.*, and is shown to be faster than traditional component labelling algorithms [10]. Each component is labelled using a contour tracing technique; as polygons are fully determined by vectors, so too are components fully determined by their respective contours, and as such all component pixels can be found. High-order objects such as characters, textlines, and text regions need to be classified in order to effectively perform a task such as DAR [11], but the application of the technique to blotch detection in film sequences appears to be novel.

A frame is processed in the same manner as when it was originally scanned – line by line, left to right from top to bottom. When a contour (either internal or external) is encountered, the tracing procedure is then applied [12] and all pixels along the contour are assigned a label, e.g., L . The contour is then traced back to its starting point, and scanning resumes. Later, when L is revisited, any neighbouring black (object) pixels are also assigned the label L . Thus, while only a single pass over the entire image is required, contour points are visited more than once due to the aforementioned contour tracing procedure, but no more than a constant number of times. Once a label index has been assigned to a pixel, its value will remain unchanged. The operations can be broken down into four major steps (see Figure 3).

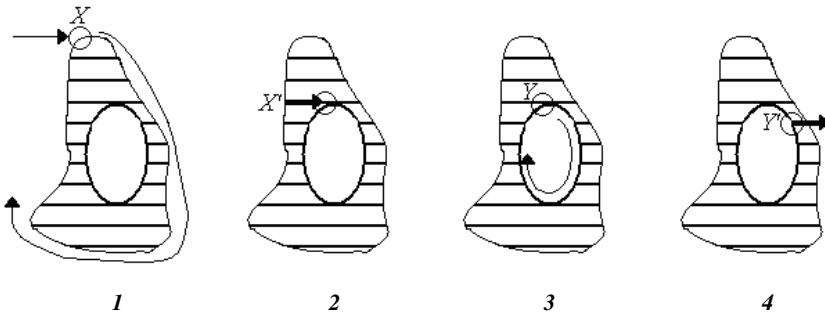


Fig. 3. The four steps of contour tracing, 1 to 4. Each line represents the current line being scanned, and the arrows show the direction of that scan.

In Step 1 when an external contour point X is encountered for the first time, a complete trace of the contour is performed until a return to X . A label is then assigned to X and all points along that contour. Step 2 shows what occurs when a labelled external contour point X' is encountered. The scan line is followed to find all subsequent black pixels (if they exist) and all are then assigned the same label. In Step 3, if an internal contour point Y is encountered for the first time, it is assigned the same label as the external contour of the same component. The internal contour containing Y is then traced, and its contour points are all assigned the same label. Finally, Step 4 dictates that when a labelled internal contour point, Y' , is encountered, all subsequent black pixels along the same scan line (if they exist) are then assigned the same label. Two separate passes of this algorithm are required, one to generate a set of potential dirt blotch candidates, the other, a set of potential sparkle. The four steps are then repeated until all components in the image have been labelled.

3 Blotch Detection

3.1 Single Frame

Once the two sets of possible blotches (one for dirt, the other for sparkle) have been generated for a given frame by application of local thresholding and contour tracing, the process of isolating genuine blotches can begin. Figure 4 shows an example of a subsection of a frame that has been thresholded to begin detection of dark blotches, for dirt.

The vast majority of object pixels will, of course, be genuine, and not at all representative of defective blotches. In order to separate and reduce false positives from potentially genuine candidates, a set of criteria has been devised.

- According to the relevant literature (and corroborating by data generated from this research to date), blotches are rarely larger than $40 * 40$ pixels, at standard definition resolutions [13]
- Blotches are isolated areas of discontinuous pixel intensities
- As previously mentioned, blotches are single frame aberrations, not occurring at the same spatial coordinates over more time

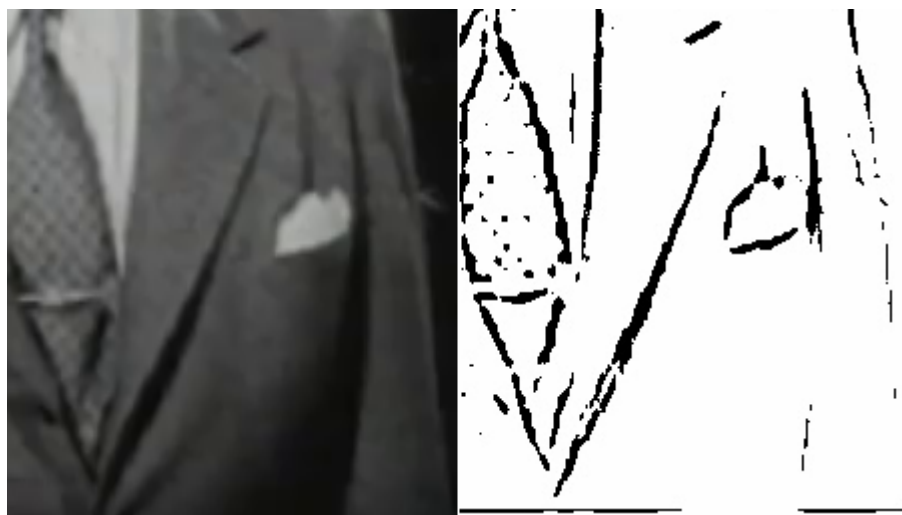


Fig. 4. On the left, close up of a frame, on the right, the locally thresholded equivalent. The adaptive threshold algorithm generates object pixels along areas of change in contrast – usually edges.

The first two criteria allow for great reduction in the number of possible blotches when using a single frame (larger values than $40 * 40$ pixels may certainly be used, but the resultant set of single frame blotch candidates will necessarily be larger.) To further reduce the number of possible blotches, an approach similar to the Simplified Ranked Order Difference (SROD) detector [14] is used. A customly laid-out neighbourhood of n approximately equidistant points just outside the external contour of the blotch under investigation are selected (see Figure 5). Standard morphological neighbourhoods are impossible to use given the irregular shape of the blotches. The selected values are then compared to the mean grayscale value of the blotch as it appears in the original frame. If greater than half of the values selected are beyond a ‘similarity of intensity’ threshold, the blotch is marked; if less, the blotch is likely to be a false positive. (If the value of n chosen is even and the result an even split, the average of all n values is compared to the threshold.) Of the 13 film sequences (from 3 different film stocks of similar age) tested to date, the application of the first two criteria checks results in a reduction of the dirt and sparkle ‘per-frame’ supersets by an average of 53%, with less than 2 blotch detection failures (from an average of 16 actual blotches) per frame. In addition, these undetected blotches could be considered minor in terms of size, and also in terms of visual discontinuity; when compared with the surrounding pixel intensities, the unfound blotches usually measured less than half of the difference of their grayscale mean and their other dirt/sparkle set equivalents in the same frame. For the remaining blotches, their classification may be further determined by applying the third criteria; exploration of a given blotch’s temporal neighbour, in a preceding and/or following frame.

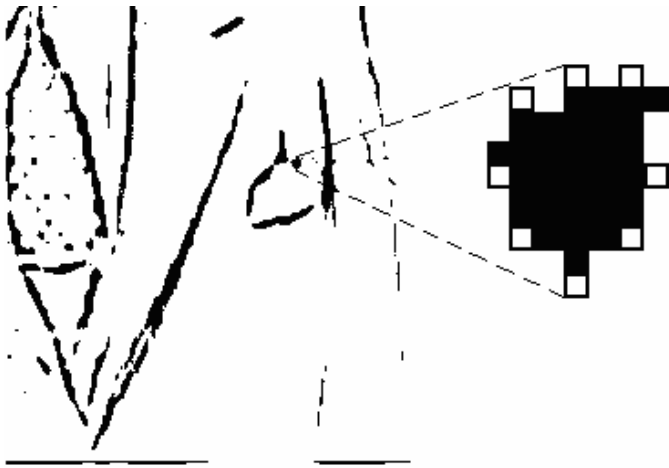


Fig. 5. Close up of a blotch with a neighbourhood of $n = 8$ (represented by the blank squares). The points are calculated by radiating out from the approximate centre of the blotch until the border has been reached, and then incremented by one.

3.2 Multiple Neighbouring Frames

In order to exploit the single frame, non-recurrence property of dirt and sparkle blotches, it is important to note that the digital frames being used must be representative of the original frame rate that the film was shot at, usually 24 progressive frames a second [15]. Altered frame rates of scanned film material can result in analogue single frame blotches becoming multiple frame digital counterparts (frame rate alteration is usually performed for television broadcast, and varies on the location depending on PAL/NTSC etc). The aim of this part of the process is to



Fig. 6. The thresholded image on the left is the first frame in a given film sequence, the image on the right, its immediate temporal neighbour. Even purely visual inspection reveals the remarkable degree of similarity between the images. The rate of motion is not high in this sequence, however, and thus is not to be considered indicative of all film sequences.

further reduce the remaining potential sets of dirt and sparkle blotches, by comparing the sets with their nearest neighbours. Given the standard twenty four frames-per-second shooting speed of motion picture cameras, consider Figure 6.

A ‘temporal location radius’, based on a given blotch’s Euclidean distance to its temporal neighbours, is then applied to each of the inter-frame blotches, *i.e.*, a permitted radius for the movement of a selected blotch based on the rate of change detected between frames, similar to instance based machine learning algorithms such as *k*-nearest neighbour [16]. A simple frame subtraction of the frame under investigation, f , from its neighbours $f + 1$ and $f - 1$ respectively can give a false impression of the rate of change due to film grain, minor illumination changes etc.; however, using the thresholded counterparts of f , $f - 1$ and $f + 1$ yields a more accurate result. If the rate of change between frames is small or large, the radius is set accordingly. By implementing this stage of the process, the numbers of potential dirt and sparkle blotch candidates are reduced by a further 18%, on average.

4 Experimental Setup and Results

While ample standard definition test footage is widely available from online resources such as the Prelinger Archives (<http://www.archive.org/details/prelinger>), and the accessibility of high definition footage becoming more common, independent benchmark footage is difficult to obtain, perhaps because of the proprietary, often copyrighted nature of the material. Consequently, the results presented in Table 1 are generated by direct comparison of the blotches detected by the system with the user annotated versions of the same film sequences. The 13 sequences that have been annotated to date represent the beginning of a much larger library that will be compiled. Given that each annotated sequence was between 2 and 5 seconds long, at 24 frames a second, an average of 100 images per sequence was visually analysed by the first author both singly and in groups of three to generate accurate data for testing purposes. Over time, with improvements in software based on the initial findings of this work, the generation of new (and larger) test data will become less involved.

The sequences were chosen as they originate from the beginning or end of a film reel, usually the most damaged sections. All of the five chosen sequences were 30 frames long, with 28 frames examined starting at frame 2 and ending at 29, allowing each frame to have a preceding and following neighbour. Based on the above results, genuine dirt blotch detection averaged 66.8%, and genuine sparkle, 62.6%. Four of the five sequences were well illuminated scenes, possibly explaining the lower result for sparkle detection as the average pixel intensity was high. The detection of false positives for both dirt and sparkle at present still appears to be randomly distributed, with no apparent connection between the instances discovered to date. True negatives vary greatly in number depending on the frame in question and the content of the footage, effectively resulting in everything else in the image not identified as a blotch. As such, while it is possible to calculate sensitivity, specificity is meaningless. The mere presence of false positives *may* not prove problematic, however, as replacement

Table 1. Comparison of blotch detection hit-rate with user generated data, from 5 chosen film sequences.

Film Sequence	Dirt Blotches Detected	False Positives	False Negatives	Num. of Actual Dirt Blotches
1	23	5	9	32
2	54	10	39	93
3	64	8	23	97
4	38	4	13	51
5	42	9	25	67

Film Sequence	Sparkle Blotches Detected	False Positives	False Negatives	Num. of Actual Sparkle Blotches
1	45	12	23	68
2	49	17	37	86
3	41	10	27	68
4	22	7	10	32
5	39	13	25	64

of missing frame data (incorrectly perceived as missing or actual) is ultimately dependent on whatever reconstruction method(s) will be employed. In conclusion, the techniques as implemented here represent a substantial first step towards automatic blotch detection, and the results significant enough (given the traditional, time consuming alternative) to warrant further research.

5 Future Work

The work completed to date has used limited samples of film stock, most of which originates from public domain archives of the 1950s. Footage from different eras and using different frame rates would make for an interesting application. Likewise, the availability of such footage is usually in standard definition – the acquisition of high definition (2K, 4K and 8K) material to suitably test the scaling of the algorithm would be highly desirable.

The automatic generation of the offsets used at the local thresholding stage and the differences of the mean grayscale intensities, as well as the selection of the value of n for the custom, border based neighbourhood would be advantageous, as would the inclusion of a locally generated temporal location radius; currently this value is global across the frame. (In many instances, whole areas of a frame sequence remain relatively static, with motion occurring in specific spatial areas.) Other blotch characteristics based on the contour geometry generated by the algorithm from Section 1.2 could be taken into consideration; perhaps the application of several machine learning algorithms could be implemented and their results compared.

References

1. Kurreck, I.H.: Diploma-Thesis Untersuchung zur Restaurierung der Sprossenschrift von Lichtton-aufzeichnungen durch digitale Bildverarbeitung, University of Applied Sciences Wiesbaden (2000)
2. Harris, R.A.: Preservation: Why Are Films and Videos Disappearing? American Film Institute, Washington (February 1993)
3. Gullu, M.K., Urhan, O., Erturk, S.: Scratch Detection via Temporal Coherency Analysis and Removal Using Edge Priority Based Interpolation. In: IEEE International Symposium on Circuits and Systems (ISCAS 2006), pp. 4591–4594 (2006)
4. Krebs, J.: Creating the Video Future, Sound & Vision Magazine (November 2004)
5. American Cinematographer, Post Focus (May 2008)
6. Kokaram, A.C.: Advances in the detection and reconstruction of blotches in archived film and video. In: Digital Restoration of Film and Video Archives (Ref. No/049), pp. 71–76. IEEE, Los Alamitos (2001)
7. Komatsu, T., Saito, T.: Detection and Restoration of Film Blotches Using Global motion Segmentation. In: ICIP 1999, vol. III, pp. 479–483 (1999)
8. Czúni, L., Hanis, A., Kovács, L., Kránicz, B., Licsár, A., Szirányi, T., Kas, I., Kovács, G., Manno, S.: Digital Motion Picture Restoration System for Film Archives (DIMORF). SMPTE Motion Imaging Journal 113, 170–176 (2004)
9. Shapiro, L.G., Stockman, G.C.: Computer Vision. Prentice Hall, Englewood Cliffs (2000)
10. Chang, F., Chen, C., Lu, C.: A Linear Time Component Labelling Algorithm Using Contour Tracing Technique. Computer Vision and Understanding 93(2), 206–220 (2004)
11. Chang, F.: Information From Document Images: Problems and Solutions. International Journal of Document Analysis Recognition, Special Issues Document, Anal. Office Syst. 4, 46–55 (2001)
12. Haig, T.D., Attikiouzel, Y.: An Improved Algorithm for Border Following of Binary Images. In: Proceedings of IEE European Conference on Circuit Theory Design, pp. 118–122 (1989)
13. Kokaram, A.: Motion Picture Restoration, ch. 6. Springer, Heidelberg (1998)
14. Gullu, M.K., Urhan, O., Erturk, S.: Blotch Detection and Removal for Archive Film Restoration. AEU - International Journal of Electronics and Communications 62(7), 534–543 (2008)
15. Brownlow, K.: Silent Films: What Was the Right Speed? Sight and Sound, 164-167 (Summer 1980)
16. Garcia, V., Debreuve, E., Barlaud, M.: Fast K Nearest Neighbor Search Using GPU. In: Proceedings of Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1–6 (June 2008)