AUTOMATED SEGMENTATION OF TORN FRAMES USING THE GRAPH CUTS TECHNIQUE

*David Corrigan, Naomi Harte, Anil Kokaram**

Department of Electronic and Electrical Engineering,
Trinity College Dublin,
Dublin 2, Ireland.

ABSTRACT

Film Tear is a form of degradation in archived film and is the physical ripping of the film material. Tear causes displacement of a region of the degraded frame and the loss of image data along the boundary of tear. In [1], a restoration algorithm was proposed to correct the displacement in the frame introduced by the tear by estimating the global motion of the 2 regions either side of the tear. However, the algorithm depended on a user-defined segmentation to divide the frame. This paper presents a new fully-automated segmentation algorithm which divides affected frames along the tear. The algorithm employs the graph cuts optimisation technique and uses temporal intensity differences, rather than spatial gradient, to describe the boundary properties of the segmentation. Segmentations produced with the proposed algorithm agree well with the perceived correct segmentation.

Index Terms— Image Restoration, Image Motion Analysis, Image Segmentation

1. INTRODUCTION

Recent years have seen the demand for archived film increase to fill the extra broadcast channels created by the growth of digital media. However, the quality of archived footage often does not meet the demands of the digital age due to degradation of the film and, as a result, restoration of archived film has become increasingly popular. Image Processing algorithms have been proposed in the past to restore sequences affected by commonly occurring forms of degradation including film shake [2], flicker [3] and the presence of dirt & sparkle (often referred to as blotches) [4].

Film Tear is a less commonly occurring form of film degradation which is caused by the ripping of the film material. The main effect of the tearing process is to introduce a global motion to a region of an affected frame (See Fig 1). Furthermore, the tearing damages the film material which results in the destruction of image data in the region of the torn edge. In
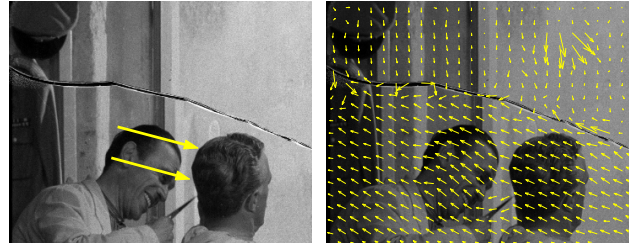
Fig. 1. Left: An example of a torn frame. The bottom of the frame has been displaced relative to the top; Right: The same frame with super-imposed motion vectors [5]. Vectors show motion relative to the previous frame in the sequence. The distribution of the vectors is clearly different either side of the tear.

[1], an algorithm was proposed to restore frames affected by tear. Its aim was to create an automatic process which would find torn frames, detect the tear boundary, estimate the global motions of the regions either side of the boundary and to compensate for the relative displacement between the regions.

The accurate segmentation of the frame into two regions is key to the success of the restoration process. It is necessary so that the global motion of each region can be robustly calculated and compensated for. The algorithm proposed for this purpose in [1] attempted to find the tear boundary and would allow for the frame to be divided into regions. However, the results from this algorithm were not sufficient to divide the frame in two as it could not robustly detect the entire path of the boundary across the image. Therefore, the frames had to be divided manually to allow the restoration to proceed.

This paper proposes a new algorithm to automatically segment torn images. Rather than detecting the boundary first, the algorithm solves the problem directly. The proposed algorithm uses the interactive graph-cuts segmentation technique presented in [6], which estimates the optimum segmentation for a two label segmentation problem. The key to the process is that it uses the tear boundary to describe the boundary properties of the segmentation (*i.e.* the segmentation boundary should coincide with the tear boundary). It also exploits the motion vector field of affected frames to automate the seg-
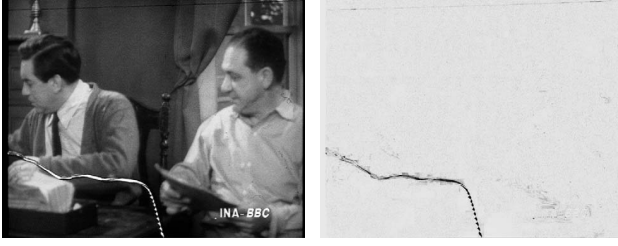
Fig. 2. The image on the right is the DFD calculated between the torn frame on the left and the previous frame in the sequence. The dark colours represent high DFD values.

mentation. The following sections of this paper introduce the proposed algorithm, the results of which are presented in Section 3. Finally, some comments on the algorithm are made.

2. TORN FRAME SEGMENTATION

The goal of the segmentation algorithm is to divide the frame into two regions. Intuitively, this problem is solved by cutting the frame along the path of the tear boundary. Although the tear boundary is characterised by high spatial gradient, it is not exclusive to the tear boundary and therefore a more unique feature is required. The chosen feature is the Displaced Frame Difference (DFD) and is derived from the observation that a tear is a temporally impulsive event and hence causes high temporal differences in the image intensity function. It is defined as a motion compensated difference as follows

$$DFD(\mathbf{x}) = |I_n(\mathbf{x}) - I_m(\mathbf{x} + \mathbf{d}(\mathbf{x}))| \quad (1)$$

where I_n and I_m are the intensity functions of the torn frame and a neighbouring frame respectively and where $\mathbf{d}$ is the motion between the frames calculated using a motion estimator such as [5].

An example of a DFD for a torn frame is shown in Fig. 2. The figure shows that the DFD forms a ridge of high values along the tear boundary. Therefore, the DFD can be considered as describing the boundary properties of the segmentation. The segmentation problem is reduced to finding the path of the ridge across the image, which is achieved using the Graph-Cuts Technique.

2.1. The Segmentation Framework

In [6], the authors present an interactive framework for segmenting an image using the Graph Cuts technique. Graph Cuts algorithms calculate the optimum binary segmentation (l) by finding the global minimum of the energy function given by

$$E(l) = \Lambda R(l) + B(l) \quad (2)$$

where

$$R(l) = \sum_{\mathbf{x}} R_{\mathbf{x}}(l(\mathbf{x})) \quad (3)$$

$$B(l) = \sum_{\mathbf{x}} \sum_{\mathbf{y} \in \mathcal{N}_s(\mathbf{x})} B_{\{\mathbf{x},\mathbf{y}\}} |l(\mathbf{x}) \neq l(\mathbf{y})| \quad (4)$$

In (2), $R(l)$ describes the region energy (likelihood) of the segmentation, with $R_{\mathbf{x}}(l(\mathbf{x}))$ being the cost of assigning a site to a given region. $B(l)$ describes the boundary (or smoothness) energy of the segmentation. The boundary energy term $B_{\{\mathbf{x},\mathbf{y}\}}$ is the energy cost of a discontinuity in the segmentation between $\mathbf{x}$ and $\mathbf{y}$, where y is in the spatial neighbourhood of $\mathbf{x}$, $\mathcal{N}_s(\mathbf{x})$. In the proposed torn frame segmentation algorithm, $B_{\{\mathbf{x},\mathbf{y}\}}$ enforces the constraint on the boundary properties discussed in Section 2 through the following equation

$$B_{\{\mathbf{x},\mathbf{y}\}} = DFD_{max} - 0.5\big(DFD(\mathbf{x}) + DFD(\mathbf{y})\big) \quad (5)$$

where DFD_{max} is the maximum DFD value. $B_{\{\mathbf{x},\mathbf{y}\}}$ is low when the mean DFD of neighbours $\mathbf{x}$ and $\mathbf{y}$ is close to the maximum and this corresponds to the case when discontinuities in the segmentation are expected in our framework.

The segmentation technique presented in [6] is initialised by allowing the user to select two subsets of pixels which must belong to a specified region. Assume that these subsets are labelled $\mathcal{L}_0$ and $\mathcal{L}_1$ such that

$$\forall \mathbf{x} \in \mathcal{L}_0, \ l(\mathbf{x}) = 0$$
$$\forall \mathbf{x} \in \mathcal{L}_1, \ l(\mathbf{x}) = 1. \quad (6)$$

For this hard constraint to be met, the following condition on the regional energies must exist

$$\forall \mathbf{x} \in \mathcal{L}_0, \ \Lambda R_{\mathbf{x}}(0) = 0 \text{ and } \Lambda R_{\mathbf{x}}(1) = \infty$$
$$\forall \mathbf{x} \in \mathcal{L}_1, \ \Lambda R_{\mathbf{x}}(0) = \infty \text{ and } \Lambda R_{\mathbf{x}}(1) = 0. \quad (7)$$

All other pixels are then assigned region energies $R_{\mathbf{x}}(l(\mathbf{x}))$ based on the probability of the pixel belonging to each region. In the proposed framework, this seeding of the segmentation is performed automatically and without the need for user intervention and is described in the following section.

2.2. Automatic Initialisation of the Segmentation

The pixel subsets $\mathcal{L}_0$ and $\mathcal{L}_1$ are estimated by exploiting the observation that tear results in each region having different global motions. As a consequence, the motion field of either side of the tear is differently distributed (Fig. 1). By performing a motion-based segmentation [7], a coarse segmentation of the frame can be estimated. This is achieved by applying a watershed segmentation to a 2-D histogram of the motion field [8] to segment the peaks of the histogram. Torn frames should contain 2 significant peaks corresponding to the global motions of both regions. Therefore, the image can be segmented by labelling blocks of pixels according to which of
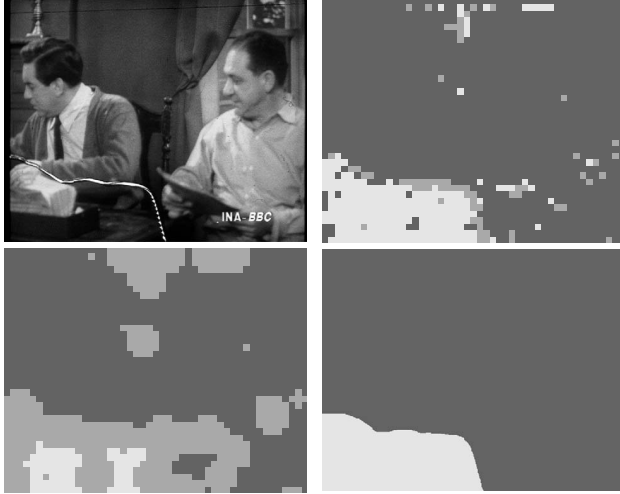
I - 558

Fig. 3. This figure shows the intermediate results of the segmentation algorithm applied to the torn frame top left; Top Right: The segmentation after the watershed segmentation is applied. The bright and dark colours represent the pixel blocks assigned to the subset of each region. The light grey colour represents blocks which are not in either subset (ambiguous); Bottom Left: Segmentation after erosion. The number of blocks in the subsets is significantly reduced; Bottom Right: The final segmentation of the frame.

Fig. 4. This figure shows the results of the segmentation (right) on two more torn frames (left). The arrows in the bottom row indicate where "shrinking" occurs during the segmentation. At the left edge of the frame the segmentation boundary takes the shorter straight path to the edge of the frame rather than the path of the tear.

the 2 peaks its motion belongs to. Pixels whose motion does not belong to either peak are marked as ambiguous. An example of such a segmentation is shown in Fig. 3.

This segmentation is still not sufficient to initialise the graph cuts segmentation according to eq. (7). Local motion or bad vectors may cause the motion at isolated pixels to be confused with the motion in the opposite region causing the pixel to be assigned to the wrong region. Furthermore, as the motion estimator used is block-based [5], the segmentation will have an undesirable blocky boundary. As a result, erosion of this rough segmentation is required to remove any wrongly labelled blocks from the subsets $\mathcal{L}_0$ and $\mathcal{L}_1$ and to allow the boundary to be computed to pixel precision. This is performed by iteratively marking blocks which have neighbouring blocks that have different labels as ambiguous. Although the erosion reduces the size of the subsets, it will ensure that all pixels in the subsets are assigned to the correct region (See Fig. 3).

2.3. Segmentation using Graph Cuts

Before the final segmentation can proceed, values need to be assigned to the region energy terms $R_{\mathbf{x}}(l(\mathbf{x}))$ for pixels not in either of the subsets. One possible method of assigning values is to compare the motion vector of a pixel to the global motion of each region using a suitable distance metric. However, for the same reason that made the erosion operation necessary
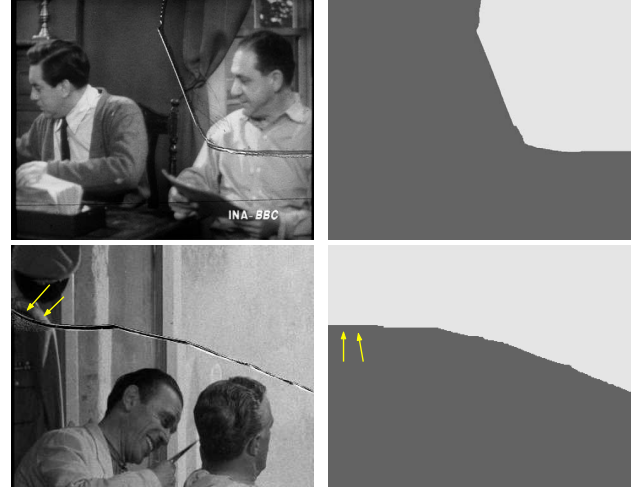
(bad motion vectors and block-based motion fields), assigning region energies in this fashion could cause the mislabelling of pixels and a blocky segmentation boundary. Therefore, values $R_{\mathbf{x}}(l(\mathbf{x}))$ for pixels not in $\mathcal{L}_0$ or $\mathcal{L}_1$ are not given. Effectively, the segmentation relies entirely on the boundary energy term given the initialisation provided by the motion based segmentation (*i.e.* $\Lambda = 0$). The optimum segmentation l is computed by using the max-flow/min-cut algorithm in [9] to minimise the energy function in (2).

3. EXPERIMENTAL RESULTS

Figures 3 and 4 (bottom right) show the results of the segmentation for 3 torn frames from 2 sequences. The segmentation maps produced by the proposed algorithm match well with the perceived correct results, as the boundary between the segmented regions generally follows the contour of the tear boundary. Some mislabelling of pixels (See Fig 4 (bottom)) occurs near the boundary of the image due to "shrinking" [6] which is a consequence of relying on the boundary energy for the segmentation. This problem could be overcome by forcing correct labelling at these sites through user-editing of the subset appropriate $\mathcal{L}_0$ or $\mathcal{L}_1$ and then re-applying the graph cuts segmentation.

After segmentation of the torn frame, the next step in the restoration process is to compensate for the displacement between the two regions. This can be achieved by estimating the global motion of each region using histograms of the vector field [1]. Finally, any damaged image data can be recovered using a missing data treatment algorithm such as [4].

Fig. 5. First Column (from left): Close-up of torn frames; Second Column: Close-up of the segmentation; Third: The torn frame with compensated displacement. The edges should now correctly align across the tear boundary; Fourth: The fully restored frame after missing data treatment. The missing data treatment algorithm used [4] also acts as a noise reducer.

4. CONCLUDING REMARKS

This paper has proposed a segmentation algorithm to divide torn frames into two regions along the tear boundary. The segmentation produced agrees well with the perceived correct segmentation. However significant motion in a frame can result in the failure of the automatic seeding process which leads to a poor segmentation. In such cases the seeding can be performed manually, allowing a good segmentation to be obtained. Work is ongoing to construct a ground truth to properly assess the accuracy of the algorithm and to find better expressions for the boundary and region energy terms.

5. REFERENCES

[1] D. Corrigan and A. Kokaram, "Automated tear treatment in degraded archived media," in *IEEE ICIP*, 2004.

[2] A. Crawford, H. Denman, F. Kelly, F. Pitie, and A. Kokaram, "Gradient based dominant motion estimation with integral projections for real time video stabilisation," in *IEEE ICIP*, 2004.

[3] F. Pitie, R. Dahyot, F. Kelly, and A. Kokaram, "A new robust technique for stabilizing brightness fluctuations in image sequences," in *the Workshop on Statistical Methods in Video Processing (ECCV)*, 2004.

[4] Anil Kokaram and Simon Godsill, "Mcmc for joint noise reduction and missing data treatment in degraded video," *IEEE trans. on Sig. Proc.*, vol. 50, 2002.

[5] Anil Kokaram, *Motion Picture Restoration: Digital Algorithms for Artefact Suppression in Degraded Motion Picture Film and Video*, chapter 2, Springer Verlag, 1998.

[6] Y. Boykov and M.-P. Jolly, "Interactive graph cuts for optimal boundary & region segmentation in n-d images," in *IEEE ICCV*, 2001.

[7] R. Coudray and B. Besserer, "Motion based segmentation using mpeg streams and watershed method," in *Int. Symposium on Visual Computing - Lecture Notes in Computer Sciences*, 2005.

[8] L. Vincent and P. Soille, "Watersheds in digital spaces: An efficient algorithm based on immersion simulations," *IEEE trans. on Pattern Analysis and Machine Intelligence (PAMI)*, vol. 13, pp. 583–598, 1991.

[9] Y. Boykov and V. Kolmogorov, "An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision," *IEEE PAMI*, vol. 26, Sept. 2004.