RESTORATION OF ARCHIVE FILM MATERIAL USING MULTI-DIMENSIONAL SOFT MORPHOLOGICAL FILTERS

Neal R. Harvey

Los Alamos National Laboratory, Los Alamos, NM, USA

ABSTRACT

A method for the optimisation of 3-D grey-scale soft morphological filters using genetic algorithms is described, which has applications in the restoration of archive film material. This method extends an existing 2-D (spatial) to the 3-D (spatio-temporal) domain, thus allows the filtering to make use of the temporal nature of the corruption and hence improve its performance.

1. INTRODUCTION

Interest has, in recent years, been increasing in the area of archive film restoration. This can no doubt be attributed to the emergence of digital television broadcasting and the increase in video sales. To meet this demand, it is becoming increasingly attractive to make archive material available to this market. Unfortunately, much archive material has been corrupted in some way and, prior to resale or broadcast, it is necessary for the material to undergo some form of restoration. Here, we describe a method for the optimisation of multi-dimensaional grey-scale soft morphological filters, with respect to specific objective image quality criteria, and their application in archive film restoration. Specifically, we address the problem known as *film dirt*.

2. THE FILM DIRT PROBLEM

Film dirt is a common problem in archive film restoration and will be familiar to anyone who has been to the cinema or viewed old film footage. Film dirt occurs when particles get caught in the film transport mechanism and damage the film, causing loss of information. This damage manifests as "blotches" of random size, shape and intensity. These blotches are non-time correlated (temporally impulsive).

Figure 1 shows an example of a region extracted from an image sequences which have been corrupted with film dirt. Stephen Marshall

University of Strathclyde, Glasgow, Scotland, UK



Figure 1: Region extracted from image sequence corrupted with film dirt

3. SOFT MORPHOLOGICAL FILTERS

Limitations to space preclude a comprehensive description of soft morphological filters. For this the interested reader is referred to [1, 2, 3, 4]. We restrict ourselves to a brief overview.

Soft morphological filters are a recently introduced class of non-linear filters [1, 2]. Their original definition was related to the class of (standard/structural) morphological filters (discrete flat morphological filters), but they have since been extended to the grey-scale (function processing) case [3]. The idea behind these filters is to relax the standard definitions of morphological filters so as to achieve robustness whilst still retaining most of the desirable properties of standard morphological filters. Whereas standard morphological operations are based on local maximum and minimum operations, with soft morphology these operations are replaced by more general weighted order statistics. In this way, the operations behave less rigidly in noisy conditions and they are more tolerant to small variations in the shapes of the objects in the filtered image.

4. OPTIMISATION OF SOFT MORPHOLOGICAL FILTERS

In [4, 5] a method was described for the optimisation of (2-D/spatial) grey-scale soft morphological filters which is able to optimise the structuring system, together with the choice of soft morphological operation. In [6] this GA optimisation technique was applied to the restoration of film material. Obviously, filtering techniques based purely in the spatial domain do not make use of the available temporal information available. The temporal characteristics of the corruption in image sequences containing film dirt (i.e. non-time correlated, temporally impulsive) provides information which may be useful in the restoration process. In this paper the 2-D (spatial) method of film dirt removal is extended to the 3-D (spatio-temporal) case in order to make use of this valuable temporal information and hence improve the restoration performance.

4.1. Soft Morphological Filter Parameters

In the search for the optimal soft morphological filter the following parameters have to be taken into consideration:

- Size and shape of structuring system's hard centre
- Size and shape of structuring system's soft boundary
- Repetition parameter
- Choice of soft morphological operations

We now describe the methods by which these parameters are incorporated into a genetic algorithm optimisation strategy. The parameters are encoded and mapped to a "chromosome".

4.1.1. Overall Structuring Function.

The outer limits to the dimensions of the overall structuring function are pre-set (i.e. the spatial, temporal and grey-scale dimensions). The optimisation process is then allowed to search for any size and shape of structuring function within this overall "envelope". If the overall dimensions of the structuring function are fixed, it may be that, for a particular structuring function, not all positions within this region are in the actual support. In order to take this into account in the GA optimisation, positions outside the structuring function's support, but within the overall search envelope, i.e. *don't care* positions, have to be distinguishable. A suitable

code, therefore, would be one which includes a unique representation for those "null" positions.

Each of these binary strings, representing an integer value for each position within the structuring function's support, can then be mapped to their appropriate positions in a string which can then be incorporated into a chromosome.

The overall structuring function is divided into two distinct regions: the hard centre and the soft boundary. Therefore there has to be some method of distinguishing these two regions incorporated into the coding scheme.

Hard Centre. A binary string, whose length is equal to the cardinality of the structuring function's overall support "envelope", is used to flag those positions within the structuring function's support which are in the hard centre. Positions in this string containing a one are positions which are in the hard centre. After forming each new individual, these hard centre flags are checked against the structuring function portion of the chromosome. If any of the positions within the structuring function portion of the chromosome are coded as being outside its support, i.e. "null" positions, then a check is made to ensure that the corresponding position within the hard centre flag string has a zero.

Soft Boundary. Any positions coded as being within the overall structuring function's support (i.e. not coded as a null position) and which do not have a one in the corresponding hard centre flag portion of the chromosome are considered to be in the soft boundary.

4.1.2. Repetition Parameter

From the definition of soft morphological operations, we know that for a structuring function, b having a support B, the repetition parameter, r, must lie in the range 1 < r < r|B|. To put it another way, the repetition parameter is related to the cardinality of the structuring function. Thus, in order to code the repetition parameter we can have a binary string, the length of which is equal to the overall size of the structuring function (i.e. the pre-set outer limits of its support). This binary string is then used to flag whether a position within the structuring function's support contributes to the repetition parameter - a one signifying that it does. To ensure consistency, after forming each new individual, a check has to be made that those positions flagged as contributing to the repetition parameter are only those positions coded as being within the structuring function's support. In this way the binary string can only code values lying within the allowable range.

4.1.3. Choice and Sequence of Soft Morphological Operations.

When considering the soft morphological operations in the context of soft morphological filter optimisation, the search

space within which the GA will operate has to be considered. Here we will limit the search to the set of fundamental (primary), secondary and tertiary soft morphological operations, i.e. to the set which includes { soft erosion, soft dilation, soft opening, soft closing, soft open-closing and soft close-opening }. Each member of this set can be defined as some combination of the fundamental soft morphological operations. Therefore, for a coding scheme to be able to encode this set of soft morphological operations, there are two basic decisions to be made;

- The set of individual soft morphological operators from which to choose.
- The maximum number of soft morphological operations in the sequence.

So, to be able to code the primary, secondary and tertiary soft morphological operations, the necessary set of soft morphological operators is { soft dilation, soft erosion} and the sequence length required is four, i.e. the longest sequence of operations will be for the tertiary operations of soft open-closing and soft close-opening, which can be defined in terms of the fundamental (primary) operations as a sequence of four separate primary operations.

In order that the GA should be able to perform optimisation over the entire search space, it is necessary to include the *do-nothing*, or *identity* operation. This is because the length of the sequence of soft morphological operations is fixed in the genetic algorithm, but it is desirable to include in the search space all the combinations of soft morphological operations from the simple soft erosion and soft dilation, through the soft close and soft open filters, to the soft openclose and soft close-open filters. Each member of the set of soft morphological operations to be considered in the GA optimisation can be coded as a bit string and these strings can then be mapped to appropriate positions in a chromosome.

It is necessary to ensure that each possible sequence of operations is unique, i.e. so that no combinations of operations in a sequences can be coded in more than one way, since some combinations of filter sequences are equivalent, e.g. soft erode, soft dilate, do-nothing, do-nothing and donothing, soft erode, do-nothing, soft dilate. This is accomplished by, after forming each new individual, checking the sequence of operations portion of the chromosome and ensuring that any do-nothing/identity operations are moved to the end of the sequence.

4.1.4. Combining the coded structuring function's hard centre and soft boundary, sequence of soft morphological operations and repetition parameter.

To form the complete chromosome, it is simply a matter of concatenating the individual strings containing the encoded structuring function, hard centre flags, sequence of soft morphological operations and repetition parameter.

The size of the search space is therefore fixed. The overall dimensions of the structuring functions - the maximum size of its support (and hence the support of the hard centre and soft boundary and the range of possible repetition parameters), the maximum grey-level values and the maximum length of soft morphological operations, together with the choice of soft morphological operations are all set beforehand. The GA will be capable of searching for any 3-D grey-level soft morphological filter which is a combination of four operations from the set {soft erode, soft dilate, donothing}, which will use a structuring function (hard centre and soft boundary) and repetition parameter chosen from all the possible variations within the overall region of support and maximum grey-level value. This search space encompasses (3-D) spatio-temporal, 2-D (purely spatial) and 1-D (purely temporal) soft morphological filters. In addition, the class of soft morphological filters encompasses several other classes of non-linear filters including standard morphological filters and rank-order filters.

4.2. Fitness Function

In order to provide each individual, representing a particular set of filter parameters with a fitness value, it is necessary to have some method of ascertaining the filter's performance, with respect to some criterion. Criteria commonly used in image processing optimisation problems usually involve some comparison of the filtered image with an "ideal" image.

Generally, in the case of film restoration, it is not possible to perform a comparison with an ideal image sequence, as such a thing does not exist. After all, if a non-corrupted version of the film exists, why bother trying to restore a corrupted version?

One method of addressing this problem is as follows. In most image sequences it is generally possible to find areas of the image which are uncorrupted. It is then possible to artificially *corrupt* this *ideal* image with particles of film dirt extracted from other similar, but corrupted, regions in the image sequence. In this way it is possible to produce the necessary *training set* which will allow the evaluation of a fitness value based on some measure of MAE.

Figure 2 shows a series of uncorrupted regions extracted from an image sequence and Fig. 3 shows the same sequence after having been artificially corrupted with film dirt.

Having a training set, (i.e. an ideal and corrupted version of the same image image sequence) enables the fitness value of an individual (i.e. a particular set of filter parameters) to be based on a comparison of the filtered image sequence (having been filtered with a filter having the parameters represented by the individual) with the ideal image sequence. The fitness of an individual is therefore determined



Figure 2: Uncorrupted regions extracted from image sequence



Figure 3: Artificially corrupted regions extracted from image sequence

as follows:

Let MAE_{max} be the maximum possible MAE for an image (for 8-bit grey-scale images MAE_{max} would be 255). Let N be the number of images in the training sequence. Let mae_i by the MAE for the i^{th} image in the filtered, corrupted sequence with respect to the i^{th} image in the ideal sequence. Let fit_i be the "interim" fitness value for the i^{th} image in the sequence and $fitness_j$ be the overall fitness of the individual j.

$$fitness_j = 0$$

for
$$i = 1$$
 to N {

$$fit_i = (1 - (mae_i / MAE_{max}))$$

$$fitness_i = (fitness_i + fit_i)$$



So, to put this in words, the "interim" fitness for an image in the sequence is a measure of how far away the filtered image is from the worse case (and hence how close it is to the ideal). The fitness value for an individual is then the average of all these interim fitness values over the whole image sequence, expressed as a percentage. A filter capable of *perfectly* restoring an image sequence would then have a fitness value of 100.

5. APPLYING THE GA OPTIMISATION PROCESS TO REAL IMAGE SEQUENCES

The GA, as described above, was run for a fixed number of generations, using the training sequence as shown in Figs. 2 and 3. Fig. 4 shows the maximum fitness at each generation during the GAs run. The best filter found is shown in Fig. 5. This filter was then applied to an entire image sequence.



Figure 4: Maximum fitness in each generation

Figure 6 shows the same image region shown in Fig. 1 extracted from the image sequence after having been filtered with the soft morphological filter found using the GA, as described.

It can be seen that the filter has done an excellent job in attenuating the corruption whilst still retaining the necessary fine details within the image.

6. CONCLUSIONS

A method has been described for the optimisation of greyscale soft morphological filters for application in the specific area of film-dirt removal from archive film material. The technique is capable of finding grey-scale soft morphological filters which have excellent performance in the removal of film-dirt corruption, whilst retaining the necessary

Hard Centre						Soft Boundary					
*	*	*	*	*		*	2	6	2	*	
*	*	*	*	*		*	*	2	*	*	
*	*	*	*	*		*	*	*	*	*	
*	*	*	*	*		*	*	2	*	*	
*	*	*	*	*		*	2	6	2	*	
*	*	*	*	*		2	6	3	6	2	
*	*	*	*	*		3	1	5	1	3	
*	*	6	*	*		5	4	*	4	5	
*	*	*	*	*		3	1	5	1	3	
*	*	*	*	*		2	6	3	6	2	
*	*	*	*	*		*	2	6	2	*	
*	*	*	*	*		*	*	2	*	*	
*	*	*	*	*		*	*	*	*	*	
*	*	*	*	*		*	*	2	*	*	
*	*	*	*	*		*	2	6	2	*	

Repetition Parameter = 39

Filter Sequence: Soft Dilation -- Soft Erosion (i.e. Soft Closing)



fine details within the original images. Examples of applying the technique to real, corrupted image sequences have been shown.

7. ACKNOWLEDGMENTS

I should like to thank the staff at BBC R & D Dept., Kingswood Warren, Surrey, England, for providing archive film material and feedback regarding the algorithm's performance and to the EPSRC for funding this work.

8. REFERENCES

- L. Koskinen, J. Astola, Y. Neuvo, "Soft Morphological Filters" *Proc. SPIE Symp. on Image Algebra and Morphological Image Processing II*, San Diego, USA, July 1992, pp. 262–270.
- [2] P. Kuosmanen, "Soft Morphological Filtering", *Ph.D. Thesis*, Dept. of Mathematical Sciences, University of Tampere, Finland, 1993.
- [3] C.C. Pu, F.Y. Shih, "Soft Mathematical Morphology: Binary and Grey-Scale", *Proc. Int. Workshop on Math*-



Figure 6: Region extracted from the filtered image sequence

ematical Morphology and its Application to Signal Processing, Barcelona, Spain, May 1993, pp. 28–33

- [4] N.R. Harvey, "New Techniques for the Design of Morphological Filters using Genetic Algorithms" *Ph.D. Thesis*, Dept. of Electronic and Electrical Engineering, University of Strathclyde, UK, 1997.
- [5] N.R. Harvey, S. Marshall, "Grey Scale Soft Morphological Filter Optimisation by Genetic Algorithms", in P. Maragos, P. Schafer, R. Butt (eds.), *Mathematical Morphology and its Applications to Image and Signal Processing*. Kluwer Academic Publishers, 1996, pp. 179– 18
- [6] N.R. Harvey, S. Marshall, "Film Restoration Using Soft Morphological Filters", Proc. 6th Int. Conf. on Image Processing and its Applications (IPIA'97), Dublin, Ireland, July 1997, pp. 279–282.

Appendix

A more detailed description of this work, together with further images, including moving image sequences, is available at the project web site ¹. This web site also contains details regarding other similar and related work in the area of film restoration.

¹ http://www.spd.eee.strath.ac.uk/harve/bbc_epsrc_film_dirt.html