CLASSIFICATION OF COMPLEX EVENT REGIONS IN OLD MOVIE MATERIAL

A. Rares, M.J.T. Reinders, J. Biemond
Information and Communication Theory Group, ITS Faculty, Delft University of Technology
e-mail: {A.Rares, M.J.T.Reinders, J.Biemond}@its.tudelft.nl

Keywords: image sequence restoration, segmentation, region matching, pathological motion.

Abstract

In this paper we address the problem of motion estimation failure in degraded image sequences. This failure is caused by some complex events that take place in the image. As a result, the sequence of operations that rely on the motion estimation process, such as motion compensation and motion picture restoration, will fail as well. We present a statistical analysis of these complex event areas, which indicates that it is possible to discriminate between the complex events resulted from complicated object motion and the ones resulted from image artefacts. An analysis scheme is proposed for the task of classifying the detected complex event areas.

1. Introduction

Recent developments in motion picture restoration show great improvements in noise reduction, flicker correction, blotch detection and removal etc. [1, 5]. However, there are situations where current techniques are still failing. In some of these cases, the quality of parts of the restored sequence is even reduced in the restoration process. For instance, in sequences where objects or persons perform complex motion - called here “pathological motion”. The pathological motion represents motion that cannot be easily modeled by current motion estimation techniques. It can be observed in areas with fast or irregular motion, occlusions, scene entrances and/or other specific circumstances [4]. When pathological motion occurs, the errors of the motion estimation process propagate into the motion compensation and image restoration processes. Here, restoration refers to the process of repairing the degraded areas of the image, rather than, for example, removing the motion blur (the latter can be even helpful by enhancing the impression of fast movement).

The usual procedure in case of pathological motion is to prevent the affected areas from being restored [6]. However, a more sophisticated method could be pursued, in order to try to restore these areas as well. To improve the detection scheme for pathological motion in degraded image sequences, a thorough analysis of this particular type of motion is essential. Based on this analysis, one can devise novel object tracking techniques [3] that are able to deal with such complex motions. These techniques will allow the subsequent restoration process to significantly improve its performance. In this paper, we address the problem of complex event classification. We propose an analysis scheme for identifying both pathological motion and artefacts. Section 2 identifies statistical properties of the pathological motion areas. In section 3 the analysis scheme is outlined in detail. Section 4 discusses the performance of our scheme. Finally, we draw some conclusions and give indications for future developments.

2. Statistical Properties of the Pathological Motion Areas

Estimating complex motion is still an open research subject [2, 8, 10]. For deteriorated image sequences, the problem becomes even harder [1, 5, 6], since motion estimation could then fail for two reasons: the pathological motion itself and/or the artefacts that affect the image content. Therefore, for restoration purposes, we propose that, in case of motion estimation failure, the underlying phenomenon to be called a complex event (CE). CE thus either refers to pathological motion (PM), or artefacts. It is necessary to discriminate between the two categories, so that the optimal restoration method can be chosen.

In a recent paper [4] we presented a taxonomy of pathological motion. For the work described here, we concentrated only on one type of pathological motion, namely the motion blur. The statistical analysis of motion blur and artefacts which follows shows that both of them can be identified on the basis of histogram examination.

---

1 This work is funded by the EU’s IST research and technological development programme. It is carried out within the Brava project (“Broadcast Archives Restoration Through Video Analysis”).
2.1 Motion blur

This class of pathological motion happens when an object moves so fast that it looks smeared over the background. We have performed a series of tests on areas with such motion blur, in order to quantify its characteristics. The tests consisted of inspecting the histograms in the RGB color space of the objects involved in the motion blur.

Fig. 1(a) shows an image of a boat with a fast moving rope. We have marked therein the upper half of the rope and a part of the background. Fig. 1(b) shows the same sequence only two frames later. Because the rope was moving very fast, it is completely smeared over the background. Investigation of the statistics of the blurred rope as well as the statistics of the non-blurred rope and background in the previous frame, shows that the histograms of the blurred rope (the dark thick-line in Fig. 1(c-e)) falls in between the histograms of the clear rope and background (the thin lines in Fig. 1(c-e)). This observation holds for all channels of an RGB image. Hence, motion blur PM can be detected by comparing the histogram of the segments inside CE areas with the histogram of object segments in previous frames.

Further experiments (not included here) have shown that the histogram of the blurred object does always fall between the other two, but not always exactly in the middle [4]. Depending on the degree of motion blur, the statistics are biased towards either the foreground object, or the background. The amount of blur, however, is the same for all channels, R, G, and B. Hence, the following ratios between the average colors on the three channels are virtually identical:

\[
\frac{\mu_c^r - \mu_{pl}^r}{\mu_{pl}^r - \mu_{pl}^b} = \frac{\mu_c^g - \mu_{pl}^g}{\mu_{pl}^g - \mu_{pl}^b} = \frac{\mu_c^b - \mu_{pl}^b}{\mu_{pl}^b - \mu_{pl}^g}
\]

(1)

where \(\{\mu_c^r, \mu_c^g, \mu_c^b\}\) represents the average color values of the current frame segment, and \(\{\mu_{pl}^r, \mu_{pl}^g, \mu_{pl}^b\}\), \(i\in\{1,2\}\), represents the average color values of the two segments from the previous frame.

These findings are in line with the existing work on motion blur [9], which is modeled as the application of a point spread function over an image.

2.2 Artefacts

In this paper we will consider only artefacts like scratches or blots. They are usually dark or bright (semi-transparent blots are rare), so their appearance is placed at the extremes of the color scale. In the example of Fig. 1(b), the artefacts are some dark and bright (bluish-green) blotsches near the left border of the complex event area. The histogram of the CE area shows “spikes” at the
extremes of the histograms, corresponding to those blotches.

3. An Analysis Scheme for Complex Event Areas

In order to apply restoration techniques in CE areas, we must separate PM from artefacts so that the most optimal restoration technique can be chosen for each area. Fig. 2 displays the general processing flow for the complete restoration chain.

First, CE areas are detected using a CE detector. Then, the analysis scheme separates them into PM and artefact areas. Finally, the image is restored by applying the optimal restoration method for each of the different PM and artefact areas in the image. Next, the first two steps are elaborated. The last one is left out of the scope of this paper.

3.1 The Complex Event Detector

We employ a CE detector [6] that makes use of a large temporal aperture, in order to detect complex events that last for more than one frame (which is the usual behaviour of the pathological motion, as opposed to blotches which usually last for a single frame). Additionally, this method makes use of a sequence of steps including: mean filters, phase correlation, minimum absolute frame difference and other low-level operators. Fig. 1(b) represents an example of the complex event detector’s output.

3.2 Analysis of Complex Event Areas

The CE analysis scheme is displayed in Fig. 3. The notation used here is as follows. \( I \) represents the current frame, \( CE \), the set of current complex event areas, and \( S \) the set of segments of the current frame. \( S_i^T \) represents the set of selected segments after a certain operation, \( s_i \) a segment, \( A(s_i) \) the area of segment \( s_i \), and \( T_S \) a selection threshold. \( T_{RPM}, T_{PM}, \) and \( T_A \), also represent thresholds. \( S_i^{A}, S_i^{AM}, \) and \( S_i^{U} \) represent the sets of artefact, pathological motion, and unclassified segments, respectively. \( D(s_i, s_j) \) represents the color difference between segments \( s_i \) and \( s_j \). \( C^{PM}(s_i, s_j) \) is a function that compares the first segment with a combination of the other two (which belong to the previous frame). \( rgbmin(s_i) \) and \( rgbmax(s_i) \) are functions that calculate the minimum and the maximum of the average colors on the \( R, G, \) and \( B \) channels, respectively.

The presented scheme first breaks up the CE area into segments that exhibit homogeneous color properties. To avoid overlapping object borders, our segments are constrained (by means of a relative over-segmentation) to only belong to one object. The CE analysis scheme then proceeds by classifying each segment into motion blur PM or artefact. This is first approached by looking in the previous frame for similar segments (Direct temporal matching). If a similar segment is found, then the current segment gets the same label as its counterpart. This ensures temporal consistency of the outcome of the analysis scheme, and speeds up the classification process, as will be shown in Section 4. When no similar segment is found in the previous frame, then the segment is classified based on the analysis of its histogram (\( PM \) and artefact classification). However, for computational purposes, instead of color histogram, the segments were represented by their average color throughout the entire scheme. This proved to be a reliable representative value of the segments. (remember that the segments are homogeneous in color, so our choice for the average color is justified)
Fig. 3. The scheme for separating pathological motion segments \( S^\text{PM}_t = S^\text{PM}_t \cup S''^\text{PM}_t \) from artefact segments \( (S^\text{A}_t = S^\text{A}_t \cup S''^\text{A}_t) \). Segments that cannot be labeled as pathological motion, nor as artifacts, are labeled as “unclassified” \( (S^\text{U}_t) \).

**Segmentation and Segment Selection**

Segmenting the CE areas consists basically of a sequence of smoothing filters for removing the noise, and a watershed segmentation. A relative over-segmentation is preferred to an under-segmentation, since we do not want adjacent distinct objects (or parts of objects) to be considered as one single segment. That is to say, we try to preserve edges, or else the color statistics of our segments become irrelevant. It should be noted that only segments within CE areas get processed, so the actual amount of segments that we must take into account will remain relatively small, even after an over-segmentation.

At this point, a problem arises from our input constraints. Namely, the complex event masks, on one hand, and the segmentation procedure used in our scheme, on the other hand, are independent from each other. The first one uses phase correlation, while the second one is based on watershed segmentation. Consequently, the boundaries of the CE masks are generally different from the boundaries of the watershed segments. Since we want to classify the segments which are related to the complex event areas, we have to use some rule to select those segments. We chose to select the segments that have at least \( T_r \) percent of their area inside the CE masks. A value of 80% for \( T_r \) proved to be a good threshold for all experiments. The set of selected segments, \( S^\text{S}_t \), is thus defined as:

\[
S^\text{S}_t = \left\{ s_j \mid s_j \in S_t, \frac{A(s_j \cap CE_t)}{A(S_t)} \geq T_r \right\}
\]  

(2)

where \( s_j \)’s represent the watershed segments, \( S_t \) the set of all the segments in the current frame, \( CE_t \) the union of all complex event areas in the current frame, \( A(s_j) \) the area of segment \( s_j \), and \( T_r \) the selection threshold of the relative area.

**Direct Temporal Matching**

This phase is triggered if we have already labeled segments of CE areas in the previous frame. Here, we look for temporal consistency between frames at the segment level. Each segment belonging to a CE area is matched with segments of the past CE areas. The unmatched CE segments are forwarded to the next phases.

The matching step takes into account the average color of the segments (for practical reasons), but one can also use other features, such as color histograms, color variance, or texture [3]. If we note
the average color of the current and previous segments with \( \{ \mu'_c, \mu'_p, \mu'_b \} \) and \( \{ \mu''_c, \mu''_p, \mu''_b \} \), respectively, then a match of the two is flagged if
\[
\frac{1}{3} \sum_{i=1}^{3} |\mu'_i - \mu''_i| < T_{\text{DRM}}
\]
where \( T_{\text{DRM}} \) is the threshold for the matching process (a value between 0 and 1), and the R, G, and B channels have been normalized to range [0…1].

**PM Classification**

In the PM classification process, we actually look for motion blur PM, in which case an object is smeared over some background. For example, the motion blur in Fig. 1(b) is a combination of the rope and background outlined in Fig. 1(a). So, for the unmatched CE segments that are handed to this phase, we try to find combinations of segments from the CE areas of the previous frame which fulfill the condition of ratio proportionality from Eq. 1. In practice, due to sampling/roundoff errors, the ratios are sometimes slightly different. We thus allow for a small variance:
\[
\sigma \left( \frac{\mu'_c - \mu'_p}{\mu'_p - \mu'_b} \right) < T_{\text{PM}}
\]
where \( \sigma \) represents the standard deviation, and \( T_{\text{PM}} \) a small threshold.

**Artefact Classification**

Segments that are not classified as PM are fed to the artefact classification process which essentially tests for extreme colors. A yet unclassified segment is labeled as “artefact” if
\[
\begin{align*}
&\min(\mu'_c, \mu'_p, \mu'_b) < T_A \\
\lor &\max(\mu'_c, \mu'_p, \mu'_b) > 1 - T_A
\end{align*}
\]
where \( T_A \) is a threshold for artefacts. We take the minimum/maximum from the R, G, and B channels and not another combination of the three values because some artefacts show up only in some of the channels (see Fig. 1(c-e)).

Artefacts have slightly different color characteristics in different sequences, so for the best results \( T_A \) has to be tuned for every image sequence. However, one could also fix a small universal threshold \( T_A \), at the expense of a higher number of unflagged artefacts. In general, we have to make a trade-off between the hit-rate (the amount of correctly flagged artefacts) and false alarms (the amount of non-artefact segments flagged as artefacts).

At the end, all segments that are not labeled as “pathological motion”, nor as “artefact”, are labeled as “unknown”. It is still possible that they may be artefacts, or pathological motion. However, for these ones there is no knowledge incorporated in the scheme yet, and they are subject to further research.

**4. Results and Discussion**

Fig. 4 and 5 show some results of our analysis scheme. We experimented with three different sequences, each of which was 50-70 frames long. The selected segments were classified either as “pathological motion” or “artefact”. The different classes were depicted in different colors: the original complex event areas are displayed in white (they also lie underneath the pathological motion and the artefacts); the pathological motion is displayed in gray; and the artefacts are displayed in black.
The results look quite promising, and a visual inspection throughout the sequence of results indicates that roughly 80% of the classification is correct. This represents a percentage relative to the amount of segments that were classified either as “pathological motion” or “artefact” (around 75% from the total number of selected segments; the other 25% remained unclassified). We rely on visual inspection because there is no ground truth to compare with. The incorrectly classified segments are due to various reasons such as: the lack of previous flicker correction of the input sequence, incorrect input CE masks, etc.

To estimate the complexity of every phase, we assume that the number of segments in a frame is \( N \), on the average. The complexity of the direct temporal matching is then of the order of \( O(N^2) \). The complexity of PM classification is \( O(N^3) \). Finally, the artefact classification is of the order of \( O(N^4) \). Obviously, direct temporal matching speeds up the classification process. It should be mentioned, however, that direct temporal matching may accumulate errors, whereas the other phases do not.

5. Future Work

We have presented here findings about the statistical content of objects performing pathological motion in the presence of artefacts. We have shown that, in principle, it is possible, based on these statistics, to discriminate between the two main forms of complex events, namely the pathological motion and the artefacts. For this purpose, we have devised a scheme for the classification of areas with complex events, where common motion estimators fail. In particular, the analysis scheme performs a segmentation of the frames and tries to find similarities between segments of the current frame and segments of adjacent frames. This inter-frame segment matching is done only for the complex event areas, thus avoiding heavy computation. The classification results are quite encouraging, and they give us hope for the future developments.

Nevertheless, for the final restoration task, one needs, besides the classification of different areas in the image, the real motion vectors as well. Since we worked at segment level and performed a segment matching between segments of the current frame, on one hand, and segments of the previous frame, on the other hand, we got some sort of segment-based motion vectors. However, to make thorough use of this information, we need to merge the segments into objects [7], so we can work with statistics and motion vectors at the object level. This step would turn our method into an object tracking algorithm that is also capable of identifying pathological motion and artefacts, and which provides object-based motion estimation - a mix of information that is needed in the last step of the restoration chain, the restoration itself.

Therefore, future work will include the development of segment-merging algorithms, restoration methods for areas degraded by artefacts, as final part of the restoration chain, and the introduction of shot change detection techniques in the restoration process. Further improvements of motion picture restoration are expected to be achieved with object modeling and tracking methods [3], as a qualitative step forward towards a higher level analysis of image sequences.

6. References


