Fade and Dissolve Detection in Uncompressed and Compressed Video Sequences

W.A.C. Fernando, C.N. Canagarajah, D. R. Bull

Image Communications Group, Centre for Communications Research, University of Bristol, Merchant Ventures Building, Woodland Road, Bristol BS8 1UB, United Kingdom.
Voice - +44-117-954-5198, Fax - +44-117-954-5206
Email - W.A.C.Fernando@bristol.ac.uk

Abstract

Automatic identification of special effects is a prerequisite for video indexing and intelligent video encoding. In this paper we present an algorithm for fade and dissolve scene change detection in video sequences. We use statistical features of the images to identify these special effects in uncompressed video. DC-estimation [4] is used to evaluate statistical features both in H.263 [10] and MPEG-2 [11] compressed video. Results show that these special effects can be identified accurately with the proposed scheme.

1. INTRODUCTION

Due to rapid advances in compression technology and imaging hardware, the expansion of low-cost storage media, and the explosion of the internet, the digital video “for every one” is now becoming a possible target. The demand for digital video is also increasing in areas such as video conferencing, multimedia authoring systems, education and video-on-demand systems. A major feature required in a visual information system is an efficient indexing to enable fast access to the stored data. A common and natural idea is to index the video sequences first into video shots by identifying scene changes and then to extract features. Therefore a powerful scene change detection, which can operate both in compressed and uncompressed domain, is required to allow for a complete characterisation of the video sequences.

Abrupt transitions are very easy to detect as the two frames are completely uncorrelated. But, gradual transitions are more difficult to detect as the difference between frames corresponding to two successive shots is substantially reduced. Considerable work has been reported on detecting abrupt transitions both in uncompressed and compressed video [1-5]. However, only a small effort has been directed toward gradual scene change detection [1,4,6-9]. Zabith et al [1] proposed a feature-based algorithm for detecting and classifying scene beaks. This algorithm needs to detect edges in every frame and hence it is very costly. Another limitation of this scheme is that the edge detection method does not handle rapid changes in overall scene brightness, or scenes, which are very dark or very bright. Furthermore, automatic segmentation and classification is not possible with this scheme. Alattar proposed an algorithm for detecting fade-in and fade-out by exploiting the semi-parabolic behaviour of the variance curve [8]. This algorithm can only detect fade-in and fade-out when the end frames are fixed (freezed). When the sequence has considerable motion, this algorithm fails to identify fade-in and fade-out regions.

In this paper we present an algorithm for dissolving and fading scene change detection in video sequences. We exploit statistical features of the luminance signal of the video to detect these gradual transitions. Rest of the paper is organised as follows: Section 2 presents a mathematical model for fading and dissolving operations in video production. Section 3 describes the proposed algorithm for fade and dissolve detection. Results are presented in section 4. Section 5 discusses the conclusions and future work.

2. MATHEMATICAL MODEL FOR DISSOLVING AND FADING

In video editing and production, proportions of two or more picture signals are simply added together so that the two pictures appear to merge on the output screen. Very often this process is used to move on from picture A to picture B. In this case, the proportions of the two signals are so that as the contribution of picture A changes from 100% to zero and the contribution of picture B changes from zero to 100%. This is called dissolving. When picture A is a solid colour, it is called as fade-in and when picture B is a solid colour, it is known as fade-out. Mathematically, dissolving can be expressed as follows.

\[
S_n(x,y) = \begin{cases} 
    f_n(x,y) & 0 \leq n < L_f \\
    \left[1 - \frac{n-L_f}{L_f}ight]f_n(x,y) + \frac{n-L_f}{L_f}g_n(x,y) & L_f \leq n \leq (L_f+L) \\
    g_n(x,y) & (L_f+L) < n \leq L_g 
\end{cases}
\]
Where, $S_n(x,y)$ - Resultant video signal, $f_n(x,y)$ - Picture A, $g_n(x,y)$ - Picture B, $L_1$ - Length of sequence A, $F$ - Length of dissolving sequence, $L_2$ - Length of the total sequence.

Assume the video sequences $f_n(x,y)$ and $g_n(x,y)$ are ergodic processes with mean $m_f, m_g$ and variance of $\sigma_f^2, \sigma_g^2$ respectively. Let, $m_{s,n}$ is the mean of the resultant video sequence ($S_n$) and $\sigma_{s,n}^2$ is the variance of the resultant video sequence. Equation (2) and (3) show the behaviour of mean and variance of the dissolved sequence.

$$m_{s,n} = E[S_n(x,y)]$$

$$\sigma_{s,n}^2 = E[S_n^2(x,y)] - E[S_n(x,y)]^2$$

where,

$$\phi = \frac{\sigma_f^2 + \sigma_g^2}{F^2} - \frac{2L_1(\sigma_f^2 + \sigma_g^2)}{F^3} - \frac{2\sigma_f^2 L_1}{F^4}$$

Similar arguments can be followed for fade-in and fade-out by simplifying the above equations. For example, fade-in can be described with the initiative function as shown in Equation (4).

$$S_n(x,y) = \left\{ \begin{array}{ll} f_n(x,y) & 0 \leq n < L_1 \\ \frac{L_1}{F} \left[ (n-L_1) \right] g_n(x,y) & L_1 \leq n \leq (L_1 + F) \\ \frac{L_1}{F} (L_1 + F) < n \leq L_2 \end{array} \right.$$
**Fade-in** Detecting a frame with zero variance followed by a sequence of continuos region, which is identified by the threshold $T_{di}$.  

**Fade-out** Detecting a sequence of continuos region, which is identified by the threshold $T_{do}$ followed by a frame with zero variance.  

**Dissolving** Detecting a sequence of continuos region, which is identified by the threshold $T_{ds}$.  

### 3.2 Detection of Fading and Dissolving in Compressed Video

Algorithms, which are proposed for uncompressed video, can still be applied in compressed domain by considering DC-sequences [4] of each frame. In compressed domain, it is considered 8x8 macroblocks (MBs) and all DC values of each MB are calculated using DC-sequence scheme [4]. Using this, mean of each frame can be evaluated approximately. However, variance may be slightly distorted due to the assumption of all the pixels have the same value in a particular MB.

### 4. RESULTS

#### 4.1 Fading

Figure 1 and Figure 2 show the mean and variance respectively for the first test video sequence. This test sequence contains one fade-in and one fade-out gradual transitions as shown in Figures 1-2. Figure 3 shows the absolute values of the differentiation in ratio between the second derivative of the variance curve to the first derivative of the mean curve. There are two regions (Figure 3), which identifies by the algorithm with the threshold $T_{di}=2$, right after the 31st frame and 111th frame. Considering the variance of the sequence, we can distinguish fade-in and fade-out regions as discussed in section 3. Therefore, fade-in region and fade-out regions are identified from 31st frame to 60th frame and from 111th frame to 150th frame respectively. 

We considered a video sequence of 2200 frames to test the proposed algorithm. This sequence contains several other special effects such as wiping, panning and dissolving. Table 1 shows the summarised results with the proposed algorithm. Results show that the algorithm is capable of detecting all fade regions accurately even when the video sequence contains other special effects. Therefore, the proposed algorithm can be used in uncompressed video to detect fade regions with a high reliability.

### 4.2 Dissolving

Figure 4 and 5 show the behaviour of mean and variance of the second test sequence considered. Figure 6 shows the absolute values of the differentiation in ratio between the second derivative of the variance curve to the first derivative of the mean curve. Thus, dissolve regions can easily be identified with the threshold $T_{ds}=2$. Hence, the dissolve regions can be identified as frame 31-60 and 121-180. We considered 1500 frames sequence with this algorithm and Table 2 shows the summarised results.

### 4.3 Fading and Dissolving in Compressed Video

Table 3 and table 4 show the results for fading and dissolving with H.263 [10] and MPEG-2 [11] compressed video. Results are slightly distorted, as variance calculation is not accurate with DC-estimation. But results are still encouraging.
Table 3: Fade region identification : compressed video with the same sequence in 4.1

<table>
<thead>
<tr>
<th>Actual fade region</th>
<th>Detected region (H.263)</th>
<th>Detected region (MPEG-2)</th>
<th>Nature of the region</th>
</tr>
</thead>
<tbody>
<tr>
<td>31-60</td>
<td>31-60</td>
<td>31-60</td>
<td>fade-out</td>
</tr>
<tr>
<td>111-150</td>
<td>112-150</td>
<td>112-150</td>
<td>fade-out</td>
</tr>
<tr>
<td>248-303</td>
<td>251-303</td>
<td>250-303</td>
<td>fade-out</td>
</tr>
<tr>
<td>576-624</td>
<td>576-627</td>
<td>576-627</td>
<td>fade-in</td>
</tr>
<tr>
<td>754-778</td>
<td>755-778</td>
<td>754-776</td>
<td>fade-out</td>
</tr>
<tr>
<td>944-986</td>
<td>944-987</td>
<td>944-987</td>
<td>fade-in</td>
</tr>
<tr>
<td>1102-1167</td>
<td>1102-1169</td>
<td>1102-1169</td>
<td>fade-in</td>
</tr>
<tr>
<td>1365-1420</td>
<td>1367-1420</td>
<td>1367-1420</td>
<td>fade-out</td>
</tr>
<tr>
<td>1500-1551</td>
<td>1500-1552</td>
<td>1500-1552</td>
<td>fade-in</td>
</tr>
<tr>
<td>1620-1680</td>
<td>1620-1682</td>
<td>1620-1681</td>
<td>fade-in</td>
</tr>
<tr>
<td>1760-1840</td>
<td>1762-1840</td>
<td>1762-1840</td>
<td>fade-out</td>
</tr>
</tbody>
</table>

Table 4: Dissolve region identification : compressed video with the same sequence in 4.2

<table>
<thead>
<tr>
<th>Actual dissolve region</th>
<th>Detected dissolve region (H.263)</th>
<th>Detected dissolve region (MPEG-2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>31-60</td>
<td>32-60</td>
<td>32-60</td>
</tr>
<tr>
<td>121-180</td>
<td>123-180</td>
<td>122-180</td>
</tr>
<tr>
<td>223-280</td>
<td>221-282</td>
<td>221-281</td>
</tr>
<tr>
<td>325-385</td>
<td>324-387</td>
<td>324-387</td>
</tr>
<tr>
<td>446-496</td>
<td>447-497</td>
<td>446-497</td>
</tr>
<tr>
<td>548-604</td>
<td>549-605</td>
<td>548-605</td>
</tr>
<tr>
<td>804-868</td>
<td>804-869</td>
<td>804-868</td>
</tr>
<tr>
<td>1010-1089</td>
<td>1012-1089</td>
<td>1012-1089</td>
</tr>
<tr>
<td>1168-1232</td>
<td>1169-1233</td>
<td>1169-1233</td>
</tr>
<tr>
<td>1356-1424</td>
<td>1356-1424</td>
<td>1356-1424</td>
</tr>
</tbody>
</table>

5. CONCLUSIONS

In this paper, we presented an algorithm for fade and dissolve scene change detection in video sequences. Statistical features of each image have been used to identify these special effects. DC-estimation [4] is used to calculate mean and variance in both H.263 and MPEG-2 compressed video. Results for uncompressed video show that these special effects can be identified accurately with the proposed scheme. Compressed domain results are slightly distorted due to the limitations of variance calculation. Further work is required to evaluate compressed domain variance accurately in order to get better results with the proposed schemes.

ACKNOWLEDGEMENTS

First author would like to express his gratitude and sincere appreciation to the university of Bristol and CVCP for providing financial support for this work.

References

Figure 1: Mean of the Luminance signal (Fading Sequence)

Figure 2: Variance of the Luminance signal (Fading Sequence)

Figure 3: Absolute change in the ratio between second derivative of the variance curve to the first derivative of mean curve (Fading Sequence)

Figure 4: Mean of the Luminance signal (Dissolving Sequence)

Figure 5: Variance of the Luminance signal (Dissolving Sequence)

Figure 6: Absolute change in the ratio between second derivative of the variance curve to the first derivative of mean curve (Dissolving Sequence)