ELSEVIER

Pattern Recognition Letters 22 (2001) 1359–1369

Pattern Recognition Letters

www.elsevier.com/locate/patrec

A fuzzy theoretic approach for video segmentation using syntactic features

R.S. Jadon [a], Santanu Chaudhury [b,*], K.K. Biswas [a]

[a] *Department of Computer Science and Engineering, Indian Institute of Technology, Hauz Khas, New Delhi 110016, India*

[b] *Department of Electrical Engineering., Indian Institute of Technology, Hauz Khas, New Delhi 110016, India*

Received 23 December 1999; received in revised form 5 March 2001

Abstract

This paper is concerned with the development of a fuzzy-logic-based framework for segmentation of video sequences. We have proposed a scheme for fuzzification of the frame-to-frame property difference values using the Rayleigh distribution. The difference values have been characterized by fuzzy terms like small, significant, large, etc. These terms have been used to design fuzzy rules for detecting abrupt changes and gradual changes. Fuzzy rules have provided a mechanism for integrating evidences based on different properties. The decompositional inference strategy has been used for fuzzy reasoning over the set of fuzzy rules. Gradual changes have been further classified as fade-in, fade-out and others (including dissolves, wipes, etc). Experimental results have shown that the proposed scheme can detect changes reliably.

Keywords: Video segmentation; Video indexing; Fuzzy theoretic; Threshold selection; Video classification

1. Introduction

Partitioning of the video sequence by detecting scene changes is essential for indexing, parsing, characterization and categorization of video. Basically, there are two types of scene changes: abrupt change and gradual change. Camera breaks or abrupt changes are apparently easy to detect because the difference in image properties between consecutive frames is expected to be large. However, the amount of change can vary from one sequence to another depending upon the content of video and context of the scene change. Detection of gradual transition is more difficult because the change takes place over a period of frames. In this paper, we propose a fuzzy theoretic framework for video segmentation. The fuzzy theoretic scheme provides a mechanism for characterizing scene transitions through soft decisions. Subjectivity involved in the video partitioning process can be adequately captured by this framework.

The problem of video segmentation has received considerable attention in the recent times. Zhang et al. (1993) propose segmentation by using a threshold setting on differences of color histograms. They also deal with special effects, such as dissolving sequences, making use of a double threshold technique. Jain et al. (1995) present a

* Correspondence author. Tel.: +91-11-659-1081; fax: +91-11-696-6606.

E-mail addresses: rsjadon@cse.iitd.ernet.in (R.S. Jadon), santanuc@ee.iitd.ernet.in (S. Chaudhury), kkb@cse.iitd.ernet.in (K.K. Biswas).

PII: S0167-8655(01)00041-1

model-based method, most suited to well-established video production processes. In particular, video segmentation is formulated as a production-model-based classification problem, and segmentation error measures are also defined. Aigrain et al. (1997) propose a rule-based system for video segmentation, which makes use of temporally located information present in video frames. These rules allow for the identification of more macroscopical changes consistent with the editing practices. In the work by Ardizzonei and Cascia (1996), a technique employing neural network is used. Essentially to detect a scene cut, two successive frames are preprocessed and a pixel-by-pixel difference is computed. This difference is then processed by a multi-layer perceptron that determines if the frames belong to the same shot. Color information present in input sequences is discarded and only luminance is used for the cut detection. Zabih et al. (1995) suggest a feature-based segmentation technique for classifying scene breaks. This approach can be used for detecting dissolves as a sequence of fade-out and fade-in effects. Xiong and Lee (1998) use a step-variable-based algorithm for scene change detection. The MPEG encoded video data can be segmented directly using the technique suggested by Arman et al. (1993). Here the DCT coefficients are analyzed to find frames where camera breaks take place.

Most of these video segmentation techniques characterize scene changes using static thresholds. Due to the time varying and subjective nature of the video data, these thresholds are not expected to work reliably under all possible conditions. Also, these approaches have used different image-based attributes for measuring the change. However, a particular property may not be always adequate for identifying the same type of transitions for different video sequences. In order to overcome these difficulties, we have proposed a fuzzy theoretic framework.

The benefits of a fuzzy framework in characterizing video data are twofold. Firstly, ambiguities and uncertainties involved in the selection of thresholds are taken care of by fuzzy rules. Secondly, because all inputs are represented as fuzzy sets, the complicated process of combination of various feature differences is simplified to straightforward application of standard fuzzy implications and compositions over these fuzzy relations.

In this paper we have presented a general scheme for fuzzifying frame-to-frame property differences. We have also indicated the patterns of fuzzy rules required for classifying scene transitions using multiple features. The actual scheme implemented in this paper is hierarchical in nature. A fusion of various syntactic features has been done to detect shot boundaries in a reliable fashion. An abrupt change is detected using histogram intersection, a gradual change is detected using a combination of pixel difference and histogram intersection while a fade is detected using a combination of pixel difference, histogram intersection and edge-pixel-count. The other syntactic features like motion (Cherfaoui and Bertin, 1995), moments difference (Arman et al., 1994) have also been tested. The motion feature has not been found to be very effective for shot detection as reported by Boreczky and Rowe (1996). There are a large number of false positives due to camera motion. However, we have found that optic-flow-based features are useful for detecting gradual transitions. Experimental results show that the proposed set of features produces reliable results.

The rest of the paper is organized as follows. We briefly discuss various syntactic features used for segmentation in Section 2. In Section 3, we have presented the fuzzy theoretic framework for characterization of the scene transitions. Experimental results are reported in Section 4. Finally, Section 5 presents the conclusions.

2. Feature extraction

Selection of an appropriate feature for segmenting video data is the most critical issue. Several such features have been suggested in the literature (Zhang et al., 1993; Tanaka and Nagasaka, 1992), like histogram difference, pixel difference, optical flow, etc. All these measures have their own advantages and disadvantages, but none is general enough to account for all types of changes in the video data. Thus instead of using a single feature, we have experimented with various combinations of features. Our fuzzy theoretic framework provides an efficient scheme for combining multiple cues. In this section, we discuss characteristics of the features used in our approach.

2.1. Histogram intersection

A histogram difference value HD_i (difference between ith and $(i+1)$th frame) is computed using normalized histogram intersection as follows:

$$HD_i = 1 - (1/3n)\left[\sum_{j=1}^{n} \min\left(F_{r_j}^{i}, F_{r_j}^{i+1}\right) + \sum_{j=1}^{n} \min\left(F_{g_j}^{i}, F_{g_j}^{i+1}\right) + \sum_{j=1}^{n} \min\left(F_{b_j}^{i}, F_{b_j}^{i+1}\right)\right],$$

where n is the number of pixels in the frame, and $F_{r_j}^{i}$ is the number of pixels in the jth bin of the red plane of the ith frame. Similar terms are defined for green and blue planes.

This measure ensures that for frames which are nearly similar, HD_i turns out to be close to zero, while for dissimilar frames HD_i is closer to one.

2.2. Pixel difference

For computing pixel difference, we have used the three-dimensional Euclidean distance between corresponding pixels in neighboring frames using R, G, B values as follows:

$$DP_i(k,l) = \begin{cases} 1 & \text{if } \left\{(R_i - R_{i-1})^2 + (G_i - G_{i-1})^2 + (B_i - B_{i-1})^2\right\}^{1/2} > T, \\ 0 & \text{otherwise,} \end{cases}$$

where $DP_i(k,l)$ is defined over the domain of two-dimensional coordinates of pixels (k,l), the subscript i denotes the index of the current frame and T is a predetermined threshold computed by doing exhaustive experimentation on various types of video sequences. Since in the second phase we are fuzzifying the pixel count, this predetermined experimental threshold works correctly for most of the video sequences. Minor change in thresholds does not affect the overall results due to inherent robustness of fuzzification scheme. By counting the number of pixels which have changed from previous frame to the next, the pixel difference metric PD_i is computed as follows:

$$PD_i = \frac{\sum_{k,l=1}^{M,N} DP_i(k,l)}{M \times N},$$

where the image has been assumed to be of size $M \times N$.

2.3. Edge pixel count

For computing edge-pixel count for each frame, firstly edges are detected using the Sobel edge detection method. It is provided as a built-in function in The X-Imaging Library (XIL) on Sun Workstations. We then count the number of edge pixels. The difference between edge-pixel count among neighboring frames is then computed. This difference metric is used for detecting fade-in and fade-out transitions among gradual transitions.

3. Fuzzy characterization of scene transitions

In this section we present a fuzzy theoretic framework for segmentation of video sequences. The framework is characterized by appropriate formulation of relevant fuzzy variables, fuzzy rules and inference methods.

3.1. Fuzzy conceptualization of variables

To use fuzzy logic and fuzzy systems for problem solving, the problem must be represented in fuzzy terms. This process is called *conceptualizing in fuzzy terms*. The objective here is to represent input and output values as *linguistic variables*. A *linguistic* variable is a variable which takes fuzzy values and has a linguistic meaning.

For the purpose of video segmentation the input variable "inter-frame-difference" is fuzzified so that it can be labeled as: "negligible", "small", "significant", "large" or "huge". The values of difference metrics as obtained in the previous section are represented as these linguistic terms. For this purpose we need to select appropriate class boundaries and membership functions for each category. Boundary values must be so assigned that these tolerate variations in individual frames while still ensuring a desired level of performance. A typical plot of frequency of occurrence of difference values is shown in Fig. 1. This particular distribution exhibits a sharp peak on the left. This indicates that a large number of consecutive frames have very small HD_i/PD_i, while a very small proportion has larger difference magnitudes. This curve suggests that inter-frame difference distribution can be modeled by the *Rayleigh distribution*. The probability density function of a Rayleigh distribution curve (Haykins, 1995) shown in Fig. 2 is given by

$$F_X(x) = \begin{cases} \dfrac{x}{\alpha^2} \exp\left(-\dfrac{x^2}{2\alpha^2}\right), & x \geqslant 0, \\ 0, & x < 0. \end{cases}$$

The cumulative distribution of this function is given by

$$F_x(x) = \begin{cases} 1 - \exp\left(-\dfrac{x^2}{2\alpha^2}\right), & x \geqslant 0, \\ 0, & x < 0. \end{cases} \tag{1}$$

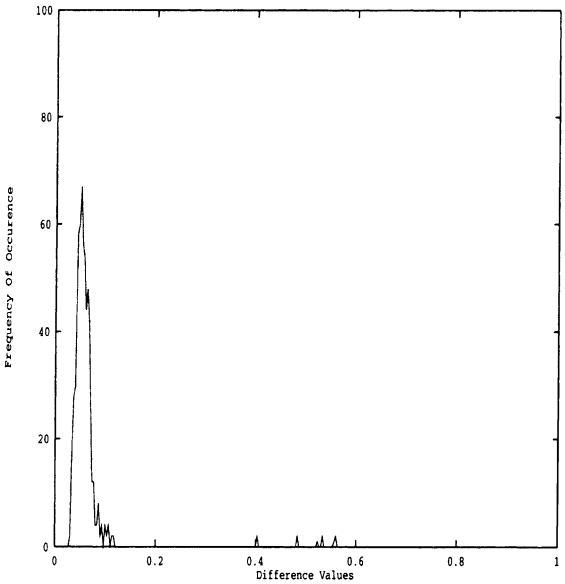

Fig. 1. Distribution of difference values.

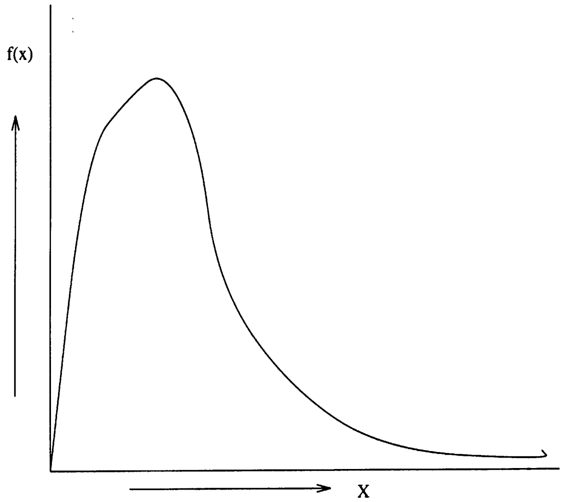

Fig. 2. Rayleigh distribution curve.

The variance of Rayleigh distribution is given by

$$\sigma = (2 - \pi/2)\alpha^2. \tag{2}$$

We have tested the appropriateness of the model by fitting Rayleigh distribution to the inter-frame difference data for nearly 300 video sequences chosen from various domains having 500–5000 frames each. The *goodness of fit* can be determined using the χ^2 test. The χ^2 values for some of the test sequences are shown in Table 1. Since the number of parameters used in estimating the expected frequencies is 1 and the total number of cells of population is 20, the degree of freedom for which chi-square value is computed $= 20 - 1 - 1 = 18$. The $\chi^2_{0.95} = 28.9$ and $\chi^2_{0.05} = 9.39$. Thus the χ^2 test of significance shows that the fit is "good" and not "so good" as to be unbelievable.

The value of the cumulative distribution function as a function of the variance is shown in Table 2. For every video sequence we compute the

Table 1
Sample χ^2 values

Sequence (Type)	χ^2 value
BBC (News)	25.7
Textile (Educational)	26.3
Bharat (Movie)	28.5
Hockey (Sports)	26.7
Ganga (Documentary)	27.3

Table 2
CDF as a function of variance

X	$F_x(X)$
σ	0.1931373
2σ	0.5761634
3σ	0.8550572
4σ	0.9677303
5σ	0.9953227
6σ	0.9995586
7σ	0.9999728
8σ	0.9999989

variance of histogram intersection and pixel difference metrics. By using the value of variance and the information in Table 2 we propose the class boundaries for different categories as shown in Table 3.

Proposed membership functions for these categories are shown in Fig. 3. This classification is based on the concept that there is a prototype or an ideal element for a class, and the degree of membership of each class is directly related to the similarity of an element to the ideal, or in other words inversely related to its distance from the ideal (prototype). Let $d(x, c)$ be the distance of an element x from the prototype element c. The fuzzy membership value of x in the class (represented by c) is

Table 3
Class boundaries for fuzzy categories

Class	Class boundaries
Negligible	0.0–2σ
Small	σ–3σ
Significant	2σ–4σ
Large	3σ–5σ
Huge	4σ–1.0

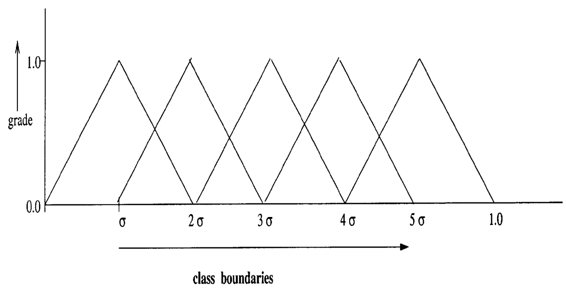

Fig. 3. Fuzzy categories.

$$M(x) = g\{d(x, c)\}. \quad (3)$$

Let l, m and h be the lower, middle and higher values for a category corresponding to difference value x. It is proposed that membership of difference value x be determined as follows:

$$M(x) = \begin{cases} (x-l)/(m-l) & \text{if } x < m, \\ 1 & \text{if } x = m, \\ (h-x)/(h-m) & \text{if } x > m, \\ 0 & \text{if } x < l \text{ or } x > h. \end{cases} \quad (4)$$

Substituting the values of l and h for different categories we obtain a membership value of x for these categories. These linguistic variables are used in designing fuzzy rules for determining the fuzzy terms: *abrupt* and *gradual* which are used to classify the scene change.

3.2. Fuzzy rules

The pattern of change between frames varies from one video sequence to another. Interpretation of this pattern is a subjective sequence dependent problem. However, analysis of the nature of the video sequences has motivated us to suggest that shot boundary is a function of two variables: quantitative difference terms between consecutive frames and similarity between these difference values in the nearest neighborhood. Thus if α is the current frame-to-frame difference, β is the previous frame-to-frame difference and γ is the next frame-to-frame difference, then the shot boundary at c, $B(c)$ can be expressed as

$$B(c) = F(\alpha, \beta, \gamma).$$

In order to accommodate variations in the quantitative value of the difference and subjective nature of the problem, we propose to model the function F in terms of fuzzy rules.

The heuristics for detecting the abrupt change is derived from its definition, i.e., *an abrupt change occurs due to a sudden and large change among neighboring frames*. This large change can occur only due to global change in image properties. Histogram intersection is an effective measure for this purpose. Thus the best choice of a feature vector used for detecting abrupt change is histogram intersection. The set of fuzzy rules for detection abrupt transition is given in Table 4.

A gradual change on the other hand can occur due to small as well as large change occurring over a long time. Thus detection of gradual transitions is difficult as compared to abrupt transitions. Here we have to take care of small as well as large difference terms. The large difference occurs due to global image properties, while small difference mostly occurs due to local changes. This scenario forced us to use the combination of feature vectors. Thus for detecting gradual transitions we use the combination of histogram intersection and pixel difference. The histogram intersection accounts for global changes and pixel difference accounts for local changes. For detecting the gradual transitions we first determine the intermediate terms H-GRADUAL (gradual change due to histogram intersection) and P-GRADUAL (gradual change due to pixel difference). The fuzzy rules for determining H-GRADUAL are given in Table 5. Similar rules are defined for P-GRADUAL (Table 6).

The fuzzy terms H-GRADUAL and P-GRADUAL obtained by applying these fuzzy rules are used to formulate the following rule for determining the gradual transition:

Rule: If $B(i)$ is H-GRADUAL and $B(i)$ is P-GRADUAL, then $B(i)$ is GRADUAL.

This combination rule gives the final membership value for gradual transition at i.

Table 4
Rules for detecting abrupt change

S. No.	Rules for abrupt change
1	If HD_i is huge and HD_{i-1} negligible and HD_{i+1} is negligible, then $B(i)$ is abrupt
2	If HD_i is huge and HD_{i-1} is negligible and HD_{i+1} is small, then $B(i)$ is abrupt
3	If HD_i is huge and HD_{i-1} is small and HD_{i+1} is negligible, then $B(i)$ is abrupt
4	If HD_i is huge and HD_{i-1} is small and HD_{i+1} is small, then $B(i)$ is abrupt
5	If HD_i is large and HD_{i-1} negligible and HD_{i+1} is negligible, then $B(i)$ is abrupt
6	If HD_i is large and HD_{i-1} is negligible and HD_{i+1} is small, then $B(i)$ is abrupt
7	If HD_i is large and HD_{i-1} is small and HD_{i+1} is negligible, then $B(i)$ is abrupt
8	If HD_i is large and HD_{i-1} is small and HD_{i+1} is small, then $B(i)$ is abrupt

Table 5
Rules for detecting H-GRADUAL change

S. No.	Rules for H-GRADUAL change
1	If HD_i is significant and HD_{i-1} is significant, then $B(i)$ is H-GRADUAL
2	If HD_i is significant and HD_{i-1} is large, then $B(i)$ is H-GRADUAL
3	If HD_i is significant and HD_{i-1} is huge, then $B(i)$ is H-GRADUAL
4	If HD_i is significant and HD_{i+1} is significant, then $B(i)$ is H-GRADUAL
5	If HD_i is significant and HD_{i+1} is large, then $B(i)$ is H-GRADUAL
6	If HD_i is significant and HD_{i+1} is huge, then $B(i)$ is H-GRADUAL
7	If HD_i is large and HD_{i-1} is significant, then $B(i)$ is H-GRADUAL
8	If HD_i is large and HD_{i-1} is large, then $B(i)$ is H-GRADUAL
9	If HD_i is large and HD_{i-1} is huge, then $B(i)$ is H-GRADUAL
10	If HD_i is large and HD_{i+1} is significant, then $B(i)$ is H-GRADUAL
11	If HD_i is large and HD_{i+1} is large, then $B(i)$ is H-GRADUAL
12	If HD_i is large and HD_{i+1} is huge, then $B(i)$ is H-GRADUAL
13	If HD_i is huge and HD_{i-1} is significant, then $B(i)$ is H-GRADUAL
14	If HD_i is huge and HD_{i-1} is large, then $B(i)$ is H-GRADUAL
15	If HD_i is huge and HD_{i-1} is huge, then $B(i)$ is H-GRADUAL
16	If HD_i is huge and HD_{i+1} is significant, then $B(i)$ is H-GRADUAL
17	If HD_i is huge and HD_{i+1} is large, then $B(i)$ is H-GRADUAL
18	If HD_i is huge and HD_{i+1} is huge, then $B(i)$ is H-GRADUAL

Table 6
Comparative segmentation results

Actual transitions		Transitions detected by non-fuzzy scheme		Transitions detected by fuzzy scheme	
Type	Frame No.	Type	Frame No.	Type	Frame No.
Gradual	29–33	Gradual	29–33	Gradual	29–33
Gradual	35–39	–	–	Gradual	36–38
Gradual	57–61	–	–	Gradual	58–61
Gradual	64–76	Gradual	64–76	Gradual	64–76
Gradual	167	–	–	Abrupt	167
Abrupt	198	Abrupt	198	Abrupt	198
–	–	Gradual	221–223	Gradual	221–223
Abrupt	230	Abrupt	230	Abrupt	230
Abrupt	274	Abrupt	274	Abrupt	274
Abrupt	306	Abrupt	306	Abrupt	306
Abrupt	1018	Abrupt	1018	Abrupt	1018
Gradual	1060–1063	Gradual	1060–1063	Gradual	1060–1063
Gradual	1104–1107	–	–	Gradual	1104–1107
Gradual	1133–1137	Gradual	1133–1137	Gradual	1133–1137
Gradual	1173–1177	Gradual	1173–1177	Gradual	1173–1177
Gradual	1195–1198	–	–	Gradual	1195–1198
Gradual	1231–1233	–	–	Gradual	1231–1233
Gradual	1273–1276	Gradual	1273–1276	Gradual	1273–1276
Gradual	1356–1359	–	–	Gradual	1356–1359
Gradual	1363–1366	Gradual	1363–1366	Gradual	1363–1366

This is basically the usage of order 1 memory. The memory of order 2 or higher was also tested, and found that effectively it has not shown any improvement for abrupt change, while some gradual transitions of smaller duration were reported as abrupt transitions due to cumulative effect. A comparison is given in Table 7.

3.2.1. Fuzzy rules for fade detection

Fade is a gradual transition with an interesting property with regard to edge-pixel-count which makes it a suitable candidate to be detected using fuzzy framework. Fade detection is based on the observation that the number of edge pixels decreases monotonically in fade-out, and increases monotonically during fade-in. Thus a fade is detected by identifying gradual transitions where this type of behavior is seen for edge-pixel-count. Edge-pixel-count is maintained for frames having gradual transition, which is then fuzzified in the same way as mentioned in previous sections for histogram intersection and pixel difference. The fuzzy terms obtained thus are used to detect fade-in and fade-out using the following heuristics:

Fade-in heuristic. If edge-pixel-count for starting frame of a gradual transition is negligible and edge-pixel-count for ending frame of this gradual transition is large then the gradual transition is of fade-in type.

Fade-out heuristic. If edge-pixel-count for starting frame of a gradual transition is large and edge-pixel-count for ending frame of this gradual transition is negligible then the gradual transition is of fade-out type.

Table 7
Comparison of order 1 and order 2 memory

Memory	No. of abrupt changes	No. of gradual changes
Order 1	6	14
Order 2	8	12

3.3. Fuzzy inferencing and rule evaluation

Since most of the rules have have more than one condition elements in the antecedent part, the *decompositional rule inference* strategy as given in (Kasabov, 1996) is used. Here a rule of k condition elements is decomposed into k implications, i.e.,

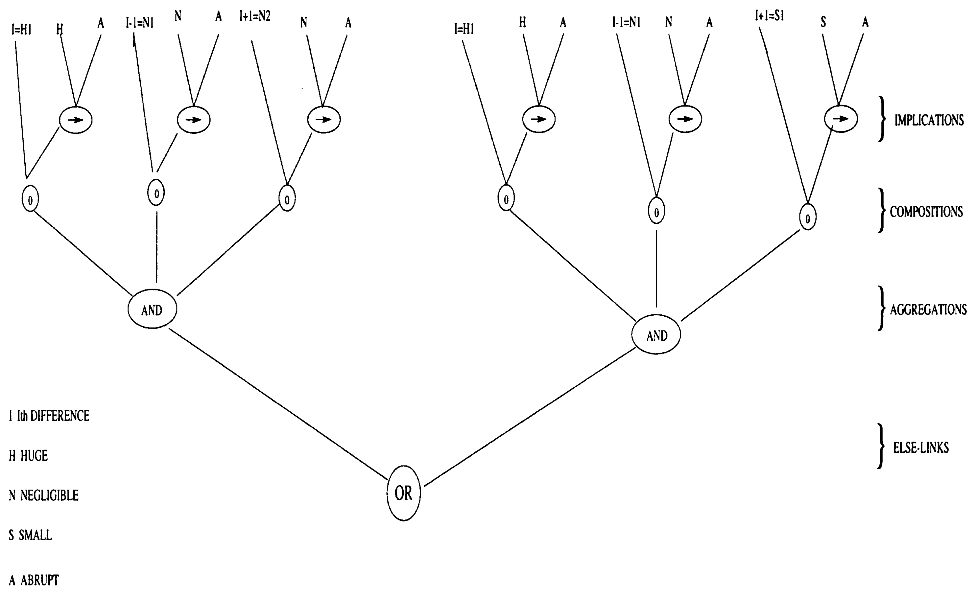

Fig. 4. Decompositional inference strategy.

$A_{ji} \to B_i$ (for $j = 1, 2, \ldots, k$). Each implication is then evaluated separately to infer a membership value for $B_{i'}$ by applying the following compositional rule:

$$B_{i'} = A_{ji'} \circ (A_{ji} \to B_i) \quad (\text{for } j = 1, 2, \ldots, k),$$

where "$\circ$" is a composition operator.

The values of $B_{i'}$ are then aggregated by the AND operator. All such rules are linked with OR–ELSE-LINKs. The implication operator used is R_c introduced by Mamdani (1977). Also the composition operator ($\circ$) used is MAX–MIN suggested by Zadeh (1965). The overall decomposition fuzzy inference method is illustrated in Fig. 4 for rules 1 and 2 of abrupt change.

4. Experimental results

The proposed scheme has been successfully implemented on Sun SPARC Ultra-1 workstation. The implementation is done in C language using XIL. We have tested the proposed scheme on a number of video sequences which include sports video, news video, documentaries, movies, etc. The corpus used to test the segmentation scheme is heterogenous in terms of effect types. It consists of some of the common cases of possible error of segmentation like sequences with fast motion, high-illumination change, different editing rhythms, etc. As compared to previous algorithms which give the crisp decision for a shot boundary, our scheme gives fuzzy decision, i.e., with every output a fuzzy membership value is attached. This membership value can be used for subsequent processing. On Sun SPARC Ultra-1 the proposed scheme took 9 frames/s.

Now, we demonstrate the performance of our scheme with reference to a particular video sequence. This is a broadcast video from a documentary film. In this video we have abrupt changes, fade-in, fade-out, dissolves and wipes. We have characterized transitions of this video using two schemes: the non-fuzzy scheme proposed in (Zhang et al., 1993) and the fuzzy one presented in this paper. Detailed classification results are shown in Table 6. The actual transitions indicated in the table have been obtained from neutral human observers. The fuzzy rules are designed such that, for every transition, the membership value is nonzero for one category only. Corresponding category has been reported in the table. As can be seen in Table 6 none of the transitions are missed in fuzzy segmentation scheme, although one false transition (classified as gradual change) is reported at 221–223. This false transition is in fact due to noise in video data. In non-fuzzy scheme many transitions are missed due to sensitivity of the static thresholds. One of the missed abrupt changes is shown in Fig. 5. This is due to the similarity of background. If we reduce the threshold such transitions can be detected but then many more false transitions are introduced. Since the fuzzy segmentation scheme is based upon fuzzy rules, this limitation is eliminated making the method less error prone.

The gradual transitions have been further analyzed for fade detection. In Table 8 we have shown how the linguistic terms based upon the number of edge pixels (for frames corresponding to gradual transitions) characterize fade-in and fade-out transitions. Results in this table show that the transition between frames 29–33 is fade-out while transition between frames 36–39 is fade-in. In general, our fade detection scheme works reliably if the length of the transition is reasonably large (typically, more than four frames).

Complete classification results for the sequence under consideration (including fade-in and fade-out results) along with membership values have been presented in Table 9. This table also shows how membership values capture prominence of the transitions. For example, the fade-out between 29 and 33 (Fig. 6) is visually more prominent as compared to fade-in between 36 and 39 (Fig. 7) and this property is reflected in the membership value of fade-out which is 0.476578 while that of fade-in is only 0.137391. Similarly membership value of abrupt change at 167 (Fig. 5) is very less as compared to 274 (Fig. 8). These results indicate that the membership values of the transitions can also be utilized for semantic characterization of the video sequences.

We have evaluated our segmentation scheme using the performance measures as suggested by Ruiloba et al. (1999). The evaluation results for the video sequence under consideration are given in Table 10.

The overall experimental results, that is, comparative performance of the fuzzy and non-fuzzy schemes for various types of video sequences is shown in Table 11. The accuracy is computed by using the formulations suggested by Ruiloba et al. (1999).

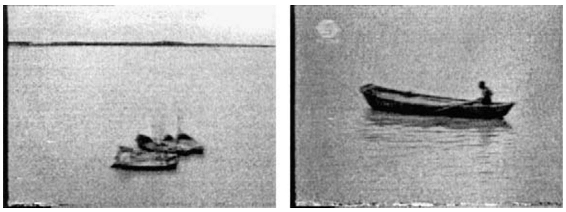

Fig. 5. Missed transition by non-fuzzy scheme (frames 167–168).

Table 8
Fade detection: Using edge-pixel-count

S. No	Frame index	Edge-pixel-count	Fuzzy value
1	29	15812.00	Large
2	30	13774.00	Significant
3	31	9821.00	Significant
4	32	6118.00	Small
5	33	3054.00	Negligible
6	36	2289.00	Negligible
6	37	2570.00	Negligible
6	38	5356.00	Small
6	39	8837.00	Significant

Table 9
Final segmentation results: Fuzzy scheme

Frame index	Type	Membership value
29–33	Fade-out	0.476578
36–38	Fade-in	0.137391
58–61	Fade-out	0.116021
64–76	Fade-in	0.616544
167	Abrupt	0.104287
198	Abrupt	0.517125
221–223	Gradual	0.189168
230	Abrupt	0.474087
274	Abrupt	0.524878
306	Abrupt	0.202835
1018	Abrupt	0.396070
1060–1063	Gradual	0.540986
1104–1107	Gradual	0.512507
1133–1137	Gradual	0.223318
150	Gradual	0.365501
1173–1177	Gradual	0.651894
1195–1198	Gradual	0.374056
1231–1233	Gradual	0.325854
1273–1276	Gradual	0.393688
1356–1359	Fade-out	0.162824
1363–1366	Fade-in	0.320281

Fig. 6. Fade-out transition (frames 29–33).

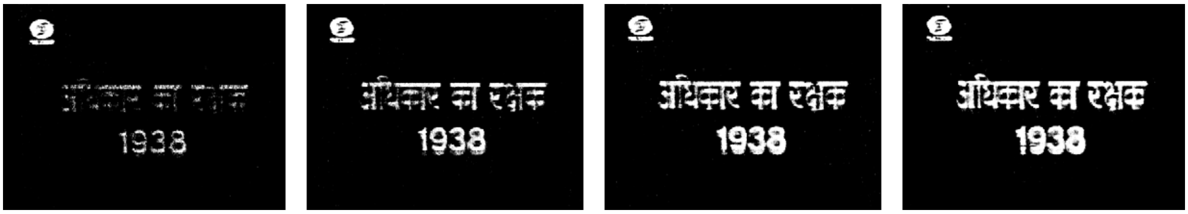

Fig. 7. Fade-in transition (frames 36–39).

Fig. 8. A prominent abrupt change (frames 274–275).

Table 10
Performance evaluation results

Quality measure	Fuzzy scheme (%)	Non-fuzzy scheme (%)
Accuracy	89.47	57.89
Precision	94.73	92.30
Recall	100	60.00
Error R	5.26	40.00

5. Conclusions

In this paper we have presented a fuzzy-logic-based video segmentation technique. This scheme takes care of subjectivity of video data while detecting shot breaks. Fuzziness involved in categorizing video data is explored for segmentation purposes. Frame-to-frame difference values are fuzzified to obtain linguistic variables. These variables appropriately quantify the amount of change occurring between consecutive frames. These variables are then used for developing fuzzy rules to detect shot breaks either as abrupt change or as gradual change. Decompositional inference strategy is used for fuzzy reasoning over fuzzy heuristic rules. Gradual transitions are further classified into fade-in, fade-out and others which includes dissolves and wipes. The final output of

Table 11
Comparative accuracy of segmentation results

Type of seq.	No. of seq.	Avg. length of seq.	Accuracy fuzzy (%)		Accuracy non-fuzzy (%)	
			Abrupt	Gradual	Abrupt	Gradual
News	110	80	85	76	56	46
Sports	96	90	87	74	54	43
Documentaries	68	220	83	71	57	41
Movies	86	300	82	68	55	35

the system is also fuzzy, i.e., with every category its membership value is also attached. This information can be further used for estimating the significance of transition and for characterizing the contents of video.

References

Aigrain, P., Joly, P., Longueville, V., 1997. Medium knowledge-based macro-segmentation of video into sequences. In: Intelligent Multimedia Information Retrieval. MIT Press, Cambridge, MA, pp. 159–173.

Ardizzonei, E., Cascia, M.L., 1996. Automatic video database indexing and retrieval. In: Multimedia Tools and Applications. Kluwer Academic Publishers, Boston, MA.

Arman, F., Depommier, R., Hsu, A., Chiu, M., 1994. Content-based browsing of video sequences. In: Proc. ACM Multimedia, San Francisco, pp. 97–103.

Arman, F., Hsu, A., Chiu, M., 1993. Image processing on compressed data for large video databases. In: Proc. 1st ACM Internat. Conf. on Multimedia, Anaheim, CA, pp. 267–272.

Boreczky, J., Rowe, L., 1996. Comparison of video shot boundary detection techniques. In: Proc. IS&T/SPIE Conf. on Storage and Retrieval for Image and Video Databases IV, Vol. SPIE 2607, pp. 170–179.

Cherfaoui, M., Bertin, C., 1995. Temporal segmentation of videos: A new approach. In: Proc. IS&T/SPIC Conf. on Digital Video Compression Algorithms and Technology, Vol. SPIE 2419.

Haykins, S., 1995. Communication Systems. Wiley Eastern Publishers, New Delhi.

Jain, R., Hampapur, A., Weymouth, T., 1995. Production model based digital video segmentation. J. Multimedia Tools Appl., pp. 9–46.

Kasabov, N.K., 1996. Foundations of Nerual Networks, Fuzzy Systems, and Knowledge Engineering. MIT press, Cambridge, MA.

Mamdani, E., 1977. Applications of fuzzy logic to approximate reasoning unsing linguistic synthesis. IEEE Trans. Comput. 26, 1182–1191.

Ruiloba, R., Joly, P., S, M.-M., Quenots, G., 1999. Towards a standard protocol for the evaluation of temporal video segmentation algorithms. In: European Workshop on Content-based Multimedia Indexing, CBMI'99.

Tanaka, Y., Nagasaka, A., 1992. Automatic video indexing and full-video search for object appearances. Visual Database Systems II, IFIP, pp. 113–127.

Xiong, W., Lee, J.C.-M., 1998. Efficient scene change detection and camera motion annotation for video classification. Comput. Vision Image Understanding 71 (2), 166–181.

Zabih, R., Miller, J., Mai, K., 1995. A feature-based algorithm for detecting and classifying scene breaks. In: Proc. Multimedia '95, San Francisco, CA, pp. 189–200.

Zadeh, L., 1965. Fuzzy sets. Inform. Control 8, 338–353.

Zhang, H., Kankanhalli, A., Smoliar, S., 1993. Automatic partitioning of full-motion video. Multimedia Syst. 1 (1), 10–28.