Video Shot-Boundary Detection based on Matrix sequence Grey model

Liu Xin[1,a], Zhang chen[1,b], Zhu hongjun[1,c]

[1]School of Software Engineering, Chongqing University of Posts and Telecommunications, Chongqing, 400065, China

[a]liuxin@cqupt.edu.cn, [b]cendra@126.com, [c]zhuhj@cqupt.edu.cn

Keywords: Grey modeling, Multi-stream information, Matrix sequence, Video Shot-Boundary Detection

Abstract. Multiple modeling is required by GM(1,1) for multi-stream sequence and there could be abnormal simulation even failed modeling sometimes. This paper analyze grey modeling for multi-stream sequence, especially, grey modeling based on matrix sequence (MGM), and research on its application in Video Shot-Boundary Detection. A method of Video Shot-Boundary Detection by MGM is proposed. The experiments show that the performance of proposed method is superior to detection method by SGM based on histogram sequence and detection method of Histogram disparity.

Introduction

Grey System Theory is an uncertain system about small samples and insufficient information, in which Grey Modeling is an effective simulation and prediction method for limited samples sequences. Grey Modeling is based on rule of grey Cause and white Effect, principle of Difference Information and Synchronous Mapping, imitating differential equation to build a model having party property of differential equation. A definition is given in the work of Professor Deng Julong [1], that is, a process of constructing an approximate differential equation model based on sequences is called as Grey Modeling, and the approximate differential equation model is called as Grey Model.

Grey Model is widely applied in various fields including image processing, economy, environmental, control field, engineering, etc. (1) Image processing, for example, image object tracking, image edge extraction, image compression and image denoising. Self-adapting template match and grey model were used by Fang to track the move of people [2]. Combing with Mean shift and GM(1,1) model, multi-feature space information is used to track video object [3]. (2) Water quality management and environmental engineering. Mahmod [4] proposed a modified Grey Model which was applied to groundwater flow analysis with limited hydrogeological data. Zhu analyzed the grey property of environment system about river based on grey theory [5]. (3) Water requirement analysis. Wang Pu et al. proposed a non-linear prediction model based on unbiased GM(1,1), and unbiased GM(1,1) and non-linear model were combined as an optimal weighted combination models to predict city water consumption in Suining [6]. Zhao et al. constructed Grey Markov prediction model to predict the city water consumption of Dalian from 2008 to 2012 [7]. (4) Other applications. Ming used a modified grey verhulst model method to Predict Ultraviolet Protection Performance of Aging B.mori Silk Fabric [8]. Chen analyzed and predicted the discharge characteristics of the lithium-ion battery based on the Grey system theory [9]. Liu predicted transportation disruptions based on an improved grey neural network model [10].

This paper research on Video Shot-Boundary Detection based on Matrix sequence grey model. The rest of this paper will illuminate as follows: Firstly, grey modeling based on matrix sequence is illustrated, and the method of Video boundary detection by grey modeling based on matrix sequence is proposed. Finally, some experiments, comparison, analysis indicates that the proposed method is effective.

Video boundary detection based on Grey modeling for multi-stream sequence

Grey modeling for matrix sequence.

Grey modeling for matrix sequence [11] and point set sequence [12] can used to simulate or predict multi-stream information. In this paper, matrix sequence is introduced here to illustrate the space relationship of video frames. According to the principle of difference information, the rule of grey Cause and Effect, Synchronous Mapping of Deng, the Grey model based on Matrix sequence (MGM) [11], which don't consider the mutual influence of elements in every matrix, can be obtained.

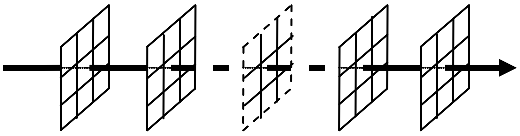

Fig. 1 The matrix sequence

Where, $ym_n = \begin{bmatrix} Xm^{(0)}(1,1,2) \\ Xm^{(0)}(1,2,2) \\ \vdots \\ Xm^{(0)}(i,j,k) \\ \vdots \\ Xm^{(0)}(m,m,n) \end{bmatrix}$, $Bm = \begin{bmatrix} -Zm^{(1)}(1,1,2) & 1 & 0 & \cdots & 0 \\ -Zm^{(1)}(1,2,2) & 0 & 1 & \cdots & 0 \\ \vdots & & & & \\ -Zm^{(1)}(i,j,k) & & & & \\ \vdots & & & 1 & \\ -Zm^{(1)}(m,m,n) & & & \cdots & 1 \end{bmatrix}$,

$Zm^{(1)}(i, j, k)$, $Xm^{(1)}(i, j, k)$ are respectively the value of $Zm^{(1)}(k)$, $Xm^{(1)}(k)$ at (i, j), ($i, j = 1, \ldots, m$, $k = 2, \ldots, n$). ym_n is a 1-column matrix, whose rows are $m^2\times(n-1)$. Bm is a matrix, whose rows are $m^2\times(n-1)$ and columns are m^2+1. The last m^2 columns in the matrix Bm are composed of 0 and 1, in which the first element in the first column is 1, and the posterior ones are 1 every other m^2; …; the k-th element in the k-th column is 1, and the posterior ones are 1 every other m^2; …; the m^2-th element in the m^2-th column is 1, and the posterior ones are 1 every other m^2; the others are all 0.

For the original matrix sequence $Xm^{(0)}$ and its mAGO generation sequence $Xm^{(1)}$, the response equation of whiten model MGM($1,1+m^2$) can be obtained:

$$\hat{X}m^{(1)}(k+1) = (Xm^{(0)}(1) - \frac{M}{a})e^{-ak} + \frac{M}{a} \quad (1)$$

$$\hat{X}m^{(0)}(k+1) = \hat{X}m^{(1)}(k+1) - \hat{X}m^{(1)}(k) \quad (2)$$

According to the two response equations, the matrix sequences could be simulated and predicted. The elements in every matrix on the prediction sequences can be obtained by $Xm^{(1)}(1)$ and the values of corresponding position in matrix M.

Video boundary detection by MGM.

The process of video boundary detection by MGM is showing as flows: Firstly, extracting the feature of video frames. To reduce computing time, some simple feature representation is used here. Some special regions at every frame is split up into 64×64 blocks. The average intensity values of every blocks is computed. These intensity values of blocks corresponds to a matrix shown as Fig.2, and then the matrix sequence as Fig.1 could be constructed. Once a new video frame comes into the detection system, the system will do grey modeling with MGM to compute *AME* (the absolute mean error) values, and the matrix sequence in the modeling process is constructed by consecutive 5 video frames including the current frame and the previous four ones. Then, another criteria, R_{AME} (regulative *AME*), will be computed at the same time as formula (4) [12]. Thresholds T and $T1$ are used to make a judgment whether a frame is the beginning of another shot. T and $T1$ relate to the intensity and size of video frames. The steps of proposed Shot-Boundary detection method are shown as Fig. 3 similar to paper [12]. In Fig. 3, $frame_i$ is the i-th frame of detected video, and i is frame number here.

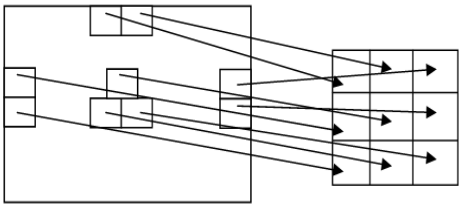

Fig. 2 Corresponding of split blocks and matrix

```
p←i, c←0, lab←0, maxA←0, j←1
While (frame_i)
    (1) obtain split blocks of frame_i, and corresponding matrix
    (2) construct matrix sequence Xm^(0)=(Xm^(0)(i-4), Xm^(0)(i-3), Xm^(0)(i-2), Xm^(0)(i-1),
Xm^(0)(i))
    (3) AME_j←grey modeling based on Xm^(0) and compute AME
    (4) (M_AME)_j←compute M_AME,
    (5) if (AME_j>T) then
                    maxA←find the local maximum
           else
          if maxA !=0 then
               lab←1
          end if
        end if
    (6) if (lab==1) then
          (6-1) c←c+1
          (6-2) if (c ==3) then
                    for q from j-3 to j step 1
                          if ((M_AME)_j<T1) then
                                frame_(p+q-4) is the beginning of another shot
                                break
                          end if
                    end for
                    c←0, lab←0, maxA←0
               end if
          end if
    (7) j←j+1
End while
```

Fig. 3 The procedure of the detection method

$$AME = \frac{1}{mn}\sum_{k=1}^{n}\sum_{i=1}^{m}\left| \mathrm{Xm}^{(0)}(k)(i) - \hat{Xm}^{(0)}(k)(i) \right| \quad (3)$$

$$R_{AME} = \frac{100 * m * AME}{\sum_{i=1}^{m}\left| \max Xm^{(0)}(k) - \min Xm^{(0)}(k) \right|} \quad (4)$$

Experiment

In this experiment, the video is CCTV news in February 6, 2015 in which include complex video effects, such as, similar background, object rapid motions, sharp illumination, strong camera flashes and dense smoke. The comparison methods are Video Shot-Boundary Detection by MGM in this paper, Video Shot-Boundary Detection by SGM in paper [12], and histogram disparity (*Dh*) between two of nearest frames I_k and I_{k+1}, shown in formula (5) in which threshold value *Th* is used to judge the cuts. The computing time of three methods are compared on equivalent calculation condition.

$$Dh = \frac{1}{256}\sum_{i=0}^{255}\left|I_{k+1}(q) - I_k(q)\right| \tag{5}$$

The following metrics is used to evaluate the performance of the proposed cut detection methods [12,14,15]:(1) Precision to measure the quality expressed as the percentage of correct detections; (2) Recall to measure the quantity expressed as the percentage of detected true cuts; (3) *F* is a measure combination of Precision and Recall.

$$Precision = \frac{\text{Number of correct detections}(CD)}{\text{Number of all detections}(AD)} \times 100 \tag{6}$$

$$Recall = \frac{\text{Number of correct detections}}{\text{Number of true cuts}} \times 100 \tag{7}$$

$$F = \frac{2 \times Precision \times Recall}{(Precision + Recall)} \times 100 \tag{8}$$

Detection results, thresholds and performance comparison of these video sets in Table 1 are demonstrated in Table 2 and Table 3 with the three methods.

Table 1 Description of the experimental video frames

Frames	**Video effects**	**True cuts (69)**	**ID**
2000-3000	object rapid motions	5	F1
18000-19000	similar background	13	F2
23000-24000	object rapid motions	14	F3
26000-27000	Illumination and camera flashes	9	F4
29000-30000	Gradual change and special effect	12	F5
32000-33000	Sharp illumination and dense smoke	16	F6

Table 2 The detection results of the three methods

ID	**DH method (*Th*=200)**			**By SGM in paper [12] (*T*=85, *T1*=11.5)**			**By MGM in this paper (*T*=1.6, *T1*=12)**		
	CD	***AD***	**Time**	***CD***	***AD***	**Time**	***CD***	***AD***	**Time**
F1	4	4	78.6s	5	5	84.9s	5	5	75.9s
F2	7	7	53s	7	7	60s	13	13	53.3s
F3	14	22	54.7s	14	15	58.7s	14	15	51.3s
F4	7	43	72s	7	20	71.9s	8	19	63s
F5	12	14	64.1s	11	12	70s	11	13	62.9s
F6	16	29	56.2s	15	16	58.9s	16	17	50.9s
total	60	119		59	75		67	82	

Table 3 The performance comparison of the three methods

ID	**DH method (*Th*=200)**			**By SGM in paper [12](*T*=85, *T1*=11.5)**			**By MGM in this paper (*T*=1.6, *T1*=12)**		
	Precision	***Recall***	***F***	***Precision***	***Recall***	***F***	***Precision***	***Recall***	***F***
F1	100	80	88.9	100	100	100	100	100	100
F2	100	53.8	70	100	53.8	70	100	100	100
F3	63.6	100	77.75	93.3	100	96.5	93.3	100	96.5
F4	16.3	77.8	26.95	35	77.8	48.3	42.1	88.9	57.1
F5	85.7	100	92.3	91.7	91.7	91.7	84.6	91.7	88
F6	84.2	100	91.4	93.8	93.8	93.8	94.1	100	96.96
total	50.4	87	63.8	78.7	85.5	82	81.7	97.1	89.3

Conclusions

On equivalent calculation condition, Video Shot-Boundary Detection by MGM in this paper can obtain better time performance. It was also observed that the performance of proposed method is superior to the other two methods.

Acknowledgement

This work is supported by the Scientific and Technological Research Program of Chongqing Municipal Education Commission under Grant no. KJ1400409; Chongqing Basic and Frontier Research Project under Grant no. cstc2015jcyjBX0090; Natural Science Foundation Project of CQ CSTC under Grant no. cstc2014jcyjA40043.

Reference

[1] J. L. Deng: Control problems of grey systems, Systems and Controls Letters (5) (1982), p. 288–294.

[2] X. Fang, J. Fang: Human Motion Tracking Based on Adaptive Template Matching and GM(1,1), in 2009 International Workshop on Intelligent Systems and Applications (2009), p. 1–4.

[3] Y. Wang, J. Zhang, L. Wu, Z. Zhou: Mean Shift Tracking Algorithm Based on Multi-Feature Space and Grey Model, Journal of Computational Information Systems Vol. 6(11) (2010), p. 3731–3739.

[4] W. E. Mahmod, K. Watanabe: Modified Grey Model and its application to groundwater flow analysis with limited hydrogeological data: a case study of the Nubian Sandstone, Kharga Oasis, Egypt, ENVIRONMENTAL MONITORING AND ASSESSMENT Vol.186(2) (2014), p. 1063–1081.

[5] C. Zhu, S. Li: Numerical Simulation of River water Pollution Using Grey Differential Model, Journal of Computers Vol. 5(9) (2010), p. 1417–1423.

[6] P. Wang, R. Chen, X. Sun, Wei Xusheng: Application of Weighted Composition Model in Urban water Consumption Forecasting, Journal of Basic Science and Engineering Vol.18 (3) (2010), p. 428–434.

[7] X. Zhao, M. Gai: Urban Water Consumption Forecasting in Dalian Based on Equal Dimensional and New Information Grey Markov Forecasting Model, Hydrology Vol.31 (1) (2011), p. 66–69, 87.

[8] J. F. Ming, Z. H. Fan, Z. G. Xie: A Modified Grey Verhulst Model Method to Predict Ultraviolet Protection Performance of Aging B.mori Silk Fabric, FIBERS AND POLYMERS Vol.14 (7) (2013), p. 1179–1183.

[9] L. Chen, B. B. Tian, W. L. Lin, et al: Analysis and prediction of the discharge characteristics of the lithium-ion battery based on the Grey system theory, IET POWER ELECTRONICS Vol. 8(12) (2015), p. 2361-2369.

[10] C. X. Liu, T. Shu, S. Chen, et al: An improved grey neural network model for predicting transportation disruptions, EXPERT SYSTEMS WITH APPLICATIONS Vol. 45(2016), p. 331-340.

[11] X. Liu, J. Dai, W. Zhou: Research on grey modeling for multi-stream information. Journal of Grey system Vol. 28 (4)(2016), p. 127-137.

[12] X. Liu, J. Dai: A Method of Video Shot-Boundary Detection based on Grey Modeling for Histogram Sequence. International Journal of Signal Processing, Image Processing and Pattern Recognition Vol. 9 (4)(2016), p. 265-280.

[13] Z. S. He, X. Liu, Y. N. Chen: Secondary-diagonal mean transformation Partial Grey Model

based on matrix series, SIMULATION MODELLING PRACTICE AND THEORY Vol. 26 (2012), p. 168–184.

[14] R. Dadashi, H. R. Kanan. AVCD-FRA: A novel solution to automatic video cut detection using fuzzy-rule-based approach. COMPUTER VISION AND IMAGE UNDERSTANDING Vol. 117 (7) (2013), p. 807-817.

[15] V. Chasanis, A. Likas, N. Galatsanos: Simultaneous detection of abrupt cuts and dissolves in videos using support vector machines. Pattern Recognition Letters Vol. 30 (1) (2009), p. 55-65.