Segmentation of Soccer Video Transitions

R. G. N. Weber Teigão
CELEPAR
Curitiba-PR, Brazil
beta_n@yahoo.com.br

J. Facon
PPGIA- PUCPR
Curitiba-PR, Brazil
facon@ppgia.pucpr.br

Abstract— In this paper, a novel approach for soccer video transition detection is proposed. This approach identifies scene cuts based on visual rhythm. A mathematical morphological lexicographic order in HSV space is used to detect scene cuts based on its color composition. Grayscale and binary morphological operators are used for false cut reduction. Experiments accomplished on soccer videos of variable quality show the promising aspects of this approach.

Keywords: MPEG Video, Visual rhythm, Scene Transition, Mathematical Morphology, HSV color space, Lexicographic order.

I. INTRODUCTION

With the increasing use of digital videos, the necessity in providing video search functionalities has transformed digital video indexing and processing in important area. In this context, detecting transitions among shots is an important step for digital video segmentation and analysis. Different kinds of transitions are present in digital videos, i.e., abrupt transitions, cuts, wipes, fades, dissolves, zooms etc... And approaches applied on uncompressed or compressed videos, and based on 2D videos or video-to-2D-image transformations to detect theses transitions are available. For instance, approaches based on dissimilarity measures [14] [4], histogram-based algorithms [11], motion-based algorithms [12], contour-based algorithms [15] or yet production-model based algorithms [6].

The visual rhythm has been presented some years ago as a new way to process videos like images. In [9], [7] one can find a complete definition of visual rhythm. On can say that visual rhythm is a single 2D image created by sub-sampling video content from row or column or yet diagonal pixels of each frame. Although a visual rhythm could only appear as a much summarized representation, most importantly, any kind of video effect is present. The discontinuities of texture and color correspond to a new event while texture and color orientation means camera manipulation and object motion [9].

By transforming a video to a single 2D image, the visual rhythm permits to directly apply various image processing techniques. We decided to use the mathematical morphology which represents one of these powerful image processing techniques. There are few approaches based on morphological tools to analyze and detect video transitions and no one based on color mathematical morphology. One can cite [5] to detect both cuts and gradual transitions or [8] for gradual change detection.

This paper presents a morphological approach to detect MPEG standard compressed soccer videos cuts from visual rhythm. No specific knowledge or mathematical modeling about soccer video is required. A color morphology based on HSV lexicographic order is employed to detect cuts in color visual rhythm. Grayscale and binary morphological operators are used to reduce false detection.

The rest of the paper is organized as follows. Section II reviews the color mathematical morphology operators and the HSV lexicographic order. Section III formally explains the video cut detection by means of visual rhythm. Experimental results over variable quality soccer videos are discussed in Section IV.

II. COLOR MATHEMATICAL MORPHOLOGY

Like binary and grayscale mathematical morphology, the color mathematical morphology is based on ordination. But differently, imposing an order on color data is not an easy task. Studies have shown that not only choosing an order but also choosing adequate color space is very important to avoid introducing color distortions.

Among the variety of color spaces and orders available in the literature, we have decided to use the lexicographic order onto HSV space proposed by [3]. There are three reasons of this choice: The lexicographical order is a complete order like dictionary ordination - The HSV color space has ability in separating luminance and chrominance information by Hue, Saturation and Value components - The lexicographic order proposed by [3] is based on a metric called Chromaticity Constant that reduces Hue and Saturation components to one value. Chromaticity Constant between two three-component vectors (h_1, s_1, v_1) and (h_2, s_2, v_2) does not use the Value component and is summarized as follows:

$$C[(h_1, s_1, v_1), (h_2, s_2, v_2)] = Sup(\mid s_1 - s_2 \mid, DistH(h_1, h_2)) \quad (1)$$

Where

$$DistH(h_1,h_2)= \begin{cases} \dfrac{|h_1-h_2|}{\Pi} & if \ |h_1-h_2| < \Pi \\ or \\ \dfrac{(2\Pi-|h_1-h_2|)}{\Pi} & if \ |h_1-h_2| \geq \Pi \end{cases}$$

Where Π= 3.14159.   (2)

The HSV lexicographical order between two three-component vectors (h_1,s_1,v_1) and (h_2,s_2,v_2) based on Chromaticity Constant is then defined as follows:

$$(h_1,s_1,v_1) < (h_2,s_2,v_2)$$
$$\Leftrightarrow C[(h_1,s_1,v_1),(h_0,s_0,v_0)] < C[(h_2,s_2,v_2),(h_0,s_0,v_0)]$$

             (3)

Where (h_0,s_0,v_0) represents the three-component vector of minimum color. In [3], the minimum color is defined as background one and it was demonstrated that this lexicographical order is suitable to process complex color images.

III.   SOCCER VIDEO CUT DETECTION APPROACH

To detail the cut detection methodology, a MPEG soccer video is used to clarify each step depicted in Figure 1. The followed methodology is applied.

A.   *Composition of soccer visual rhythm:*

From pixel diagonal of each frame (using only the DC values) the visual rhythm (2D image) is created. Figure 1-(a) illustrates the visual rhythm which summarizes 45 minutes of soccer video. It is possible to observe the complexity of the visual rhythm which depicts two textured regions: one of them, dominantly green, represents the summarization of events occurred in grass field (mainly the game). The second one corresponds to summarization of other events (public, publicity, player zoom, etc…).

The four real cuts are indicated at the upper part of image. Two of them are present in the green grass field texture and are not very discernible. Two other cuts delimit the public/publicity texture.

B.   *Transition detection:*

The transitions are detected by vertical morphological erosion based on HSV lexicographical order. In this step all vertical dominant edges are detected. Figure 1-(b) shows the transition image. One can observe that the real cuts are mixed with grass field and public/player borders;

C.   *Transition filtering:*

The previous image is converted to grayscale one and then filtered by means of vertical closing-opening alternate sequential filter [13]: this step aims to remove noise and highlight real cuts. In Figure 1-(c) we can observe that real cuts are preserved while other borders are removed in grass field texture.

D.   *Binarization:*

The previous image is binarized by means of two threshold processes, the Global Pun´s one [10] and the Local Bernsen´s one [1]: the first process one allows to preserve all vertical dominant cuts, while the second one only detects markers of principal cuts (figures 1-(d) and (e));

E.   *Filtering of cut candidates:*

This process aims to preserve complete vertical cuts and to delete interrupted edges. This filtering is performed by reconstructing the Bernsen´s threshold image from the Pun´s threshold one. Figure 1-(f) illustrates the cut filtering.

F.   *False positive Reduction:*

MPEG standard defines three different frames, I (intra frames), P (forward predicted frames) e B (bi-directionally predicted frames). No real cut can occur between two consecutive intra frames I and (I+1). The presence of transitions between two consecutive intra frames is originated from camera manipulations. This kind of transitions does not represent real cuts. The false positive reduction is performed by eliminating any detected vertical line between two consecutive intra frames.

Figure 1-(g) illustrates the result of the soccer cut detection obtained by the complete methodology.

IV.   EVALUATION AND EXPERIMENTAL RESULTS

The proposed method was tested and evaluated on six videos of soccer games with different quality (30 Fps) totalizing 540 minutes. The ground-truth scene cut location for each video was manually carried out by using the V*idsegpick* tool [2]. For instance, Figure 1-(a) indicates the four ground-truth cuts.

Precision, Recall and Error criteria, given by (4), are adopted to evaluate the scene cut detection performance of the proposed approach. These can be defined by:

$$precision = \frac{T^+}{T^+ + F^+} \ , \ recall = \frac{T^+}{T^+ + F^-} \ , \ error = \frac{F^+}{T^+ + F^-}$$

             (4)

where T^+ represents true cuts (that is real cuts in the ground-truth that are correctly detected by the approach), F^+ represents false positive cuts (not cuts that the methodology has wrongly labeled as real cuts) and F^- represents real cuts not labeled as real cuts, (real cuts in the ground-truth not detected by the methodology).

Table I presents the proposed approach evaluation. The average Precision, Recall and Error rates are 78%, 80.9% and 23.6% respectively. The common use of camera effects like zoom-ins/outs or abrupt camera

movements in soccer videos increase the complexity of such videos. Great variety of weather, lighting, color variations and quality of soccer games also represent complex factors. Considering the huge complexity of soccer videos, theses results are highly significant. Based on these numerical results, we can say that our approach is very promising.

Figure 2 illustrates four complex situations analyzed by the proposed method:

- Figure 2-(a) shows two real cuts that are not very discernable and mixed with green grass field texture. The proposed approach presets correct detection in this very complex situation. The use of HSV lexicographical order was determinant to clearly detect them;

- Figure 2-(b) depicts two real cuts not very discernable. The proposed method has only succeeded in detecting one of them, resulting in a true positive T^+ and a false negative F^-;

- Figure 2-(c) presents the visual rhythm of a daylight soccer game. The two transitions created by lighting variations are not real cuts and have been well interpreted and ignored by the proposed method. We consider that the use of color morphological operators was determinant to reject them; Figure 2-(d) presents the visual rhythm with a small brightness change or compression error that looks like a cut transition but is not a real cut. This transition has ignored by the proposed method.

The results in Table I and examples in Figure 2 show that we have to improve the correct detection and false detection results.

TABLE I.　　QUANTITATIVE PERFORMANCE EVALUATION

Soccer Video	Precision (%)	Recall (%)	Error (%)
Number 1	86.7	82.0	12.5
Number 2	82.5	80.5	17.1
Number 3	72.7	88.9	33.3
Number 4	73.2	83.0	30.4
Number 5	80.0	68.4	17.1
Number 6	72.5	82.7	31.3
Average	78.0	80.9	23.6

V.　CONCLUSIONS

A methodology to identify cut transitions in soccer games was proposed. The scheme has consisted in detecting cuts from the visual rhythm which represents a new concept that transforms a video in a single 2D image. This study shows how the use of color, grayscale and binary morphological operators can efficiently filter and detect scene transitions in complex visual rhythm. The use of color mathematical operators based on HSV lexicographical order was determinant to clearly detect

real cuts and discard false cuts created by lighting variations. By using grayscale visual rhythm instead of color one, these weak variations can not be well segmented by grayscale morphological operators.

The numerical results show that this procedure is promising and could be used to identify transitions in other kinds of video. Future work will involve detecting other types of scene transitions like fade in, fade out etc.

REFERENCES

[1] Bernsen, J., "Dynamic Thresholding of gray-level images", Proc. Eighth Int'l Conf. on Pattern Recognition, Paris, France, pp 1251-1255, 1986.

[2] Bose, P., Laganiere, R., and Whitehead, A., "Vidsegpick", 2003. available at http://iv.csit.carleton.ca/~awhitehe/vidproc/

[3] Petito, E.P. Conci, A., "Chromaticity constant: Introducing a new ordination for automated extraction of grain–size data from true colour images", 8th International Symposium on Mathematical Morphology, ISMM 2007, vol. 2, pp 63-64 , 2007

[4] Del Bimbo A. "Visual Information Retrieval", Morgan Kaufmann, 1999.

[5] Guimarães, S.J.F.; de Araujo, A.; Couprie, M.; Leite, N.J., " An approach to detect transitions based on mathematical morphology", ICIP 2003 International Conference on Image Processing, vol. III, pp. 1021-1024, 2003

[6] Hampapur A., Jain R., Weymouth T., "Production Model based Digital Video Segmentation," Multimedia Tools and Applications, vol.1, no.1, pp 9-46, 1995.

[7] Kim Hyeokman, Lee Jinho, Yang Jae-Heon, Kim Woonkyung M., Song S. Moon-Ho Song, "Visual Rhythm and Shot Verification", Multimedia Tools and Applications, 15, pp227–245, 2001

[8] Naranjo, V., Angulo, J., Albiol, A., Mossi, J. M., Albiol, A. , Gómez, S., "Gradual Transition Detection For Video Partitioning Using Morphological Operators", Image Analysis And Stereology, vol. 26, Issue 2, pp 51-61,2007

[9] Ngo C. W., Pong T. C., and Chin R. T. "Detection of gradual transitions through temporal slice analysis", Proceeding of the IEEE CVPR, pp 36–41, 1999.

[10] Pun T., "Entropic Thresholding, the new Approach", Computer Graphics and Image Processing, vol.16, pp 210-239, 1981.

[11] Sethi I. K. , Nilesh, P. "A Statistical Approach to Scene Change Detection", SPIE, vol. 2420, pp 329-338, 1995.

[12] Shahraray B., "Scene Change Detection and Content-based Sampling of Video Sequences", Digital Video Compression: Algorithms and Technologies, pp 2-13, 1995.

[13] Soille P., "Morphological Image Analysis: Principles and Applications", Springer-Verlag New York, 2003.

[14] Wang Y., Liu Z., and Huang J.-C. "Multimedia content analysis", IEEE Signal Processing Magazine, pp 12– 36, 2000.

[15] Zabih R., Miller J., Mai K., "Feature-based Algorithms for Detecting and Classifying Scene Breaks", Proc. Fourth ACM International Conf. on Multimedia, pp 189-200, 1995.
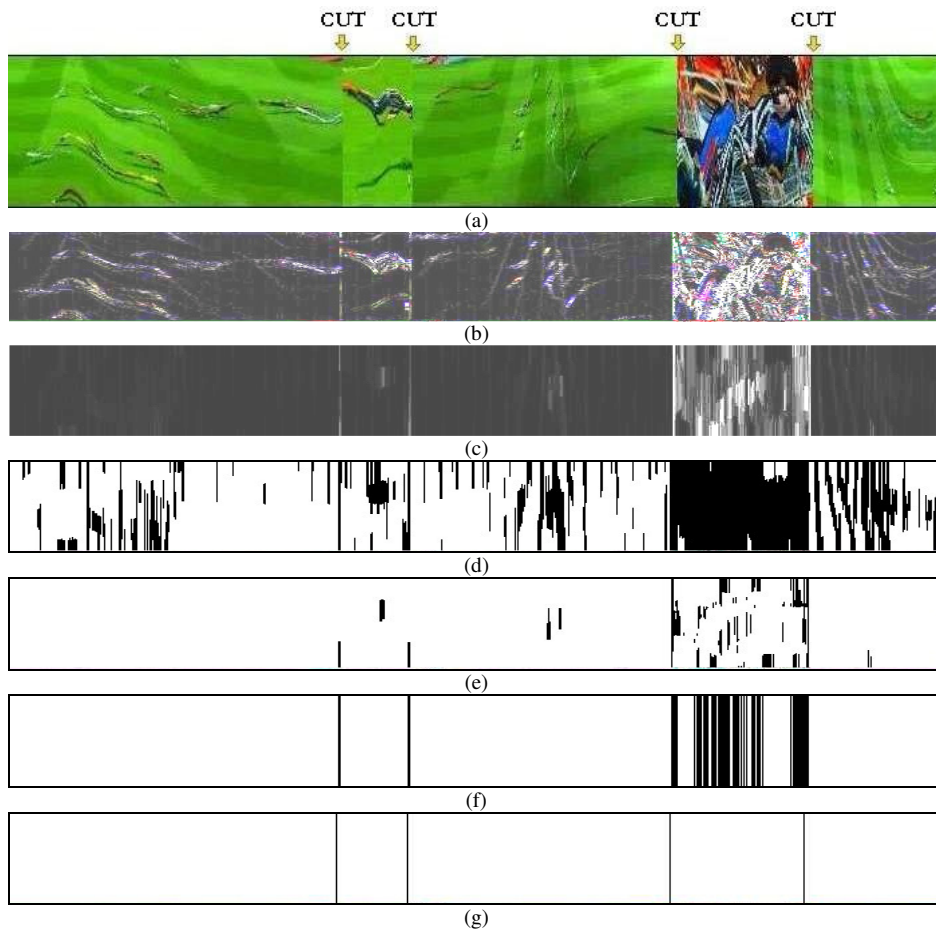
Figure 1: Cut detection: (a) visual rhythm with the four real cuts – (b) Transition detection - (c) Transition filtering – (d) Global Binarization – (e) Local Binarization – (f) Candidate filtering – (g) Final result
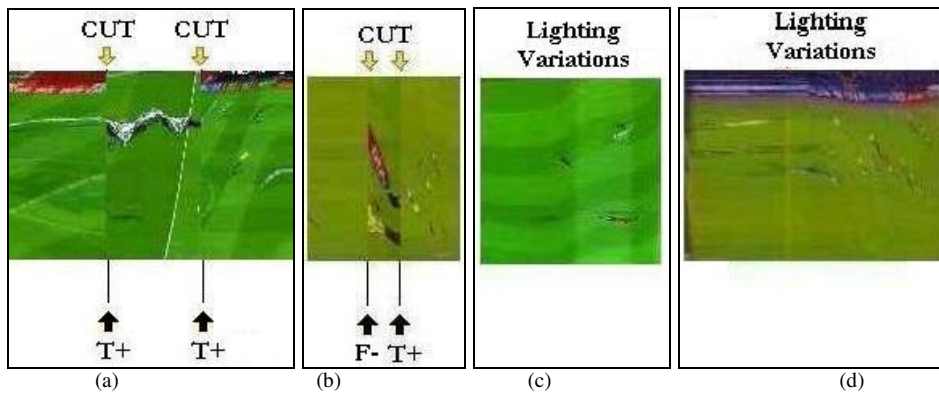

Figure 2: Complex Examples: (a) True positive T^+ - (b) True positive T^+ and False negative F^- - (c & d) Good interpretation