A Real-Time System for Detecting Indecent Videos Based on Spatiotemporal Patterns

Soonhong Jung, Junsic Youn, and Sanghoon Sull, *Senior Member*, IEEE

Abstract — *In this paper, an efficient real-time system is proposed for detecting indecent video scenes, which usually contain periodically moving skin-colored objects. Spatiotemporal motion trajectories are efficiently extracted by the simple color-based region segmentation of a spatiotemporal pattern generated from an input video. By analyzing the vertical displacement of the spatiotemporal motion trajectory along the time axis of the spatiotemporal pattern, the trajectory is then converted to a one-dimensional (1D) signal. Feature vectors computed from the discrete Fourier transform (DFT) of the 1D signal and the colors near the spatiotemporal motion trajectory are used as the inputs for a classifier used to detect adult scenes. Experimental results show improvements in true positive and false alarm rates when compared to existing methods, and significantly reduced processing times*[1].

Index Terms — Indecent video, spatiotemporal pattern, periodic motion, spatiotemporal motion trajectory.

I. Introduction

The rapid growth and popularity of social networking services have enabled users to easily upload and share videos with each other. A negative consequence of this trend is that children with smartphones can access harmful material on the Internet with little difficulty, and thus, to protect them from indecent videos or adult scenes, an efficient, low-power method for automatically detecting and filtering these scenes from their smartphones is proposed in this paper.

A typical adult scene is commonly characterized by the presence of two key features: periodic motion and human skin color. Most existing methods for detecting adult scenes are based on spatial and/or temporal features. Methods based on spatial features utilize skin color [1]-[4] and a scale invariant feature transform (SIFT) descriptor by employing a bag of visual words (BOVW) model [3]-[4]. Although the BOVW approach provides an effective image representation, its computational cost is high, even when applied only to key video frames.

Other methods utilize the medians of motion vectors in compressed bit streams to detect dominant repetitive motions [5]-[6]. However, this depends on a particular type of compression and the motion vectors in bit streams are often unreliable, especially for uniform regions within a frame. Moreover, it is difficult to distinguish object motion from background motion when using the medians of motion vectors.

Recently, approaches based on the combined use of spatial and temporal features have been proposed [7]-[10]. Lee *et al*. [7] used a skin-colored patch as a spatial feature, and the signature from video file headers and a color histogram as temporal features. Kim *et al*. [8] utilized global motion between frames, skin color, and shape matching based on Zernike moments. However, neither of these methods can reliably detect periodic motion. Behrad *et al*. [9] detected the largest skin-colored region, computed its correlation between frames, and analyzed the DFT of the correlation to detect adult scenes. However, this can be problematic if the largest region does not move periodically. Ulges *et al*. [10] combined various features such as skin color, the discrete cosine transform coefficients of local image patches based on a BOVW model, motion histograms (MHIST), and the mel-frequency cepstral coefficient (MFCC). However, when only visual cues are used, erroneous results are sometimes produced because reliably detecting periodic motion using MHIST is difficult. Further, the computational cost is high.

Other approaches that detect periodic motion are based on a spatiotemporal pattern obtained by sampling sets of pixels in each frame, and temporally accumulating them [11]-[13]. This requires the moving objects to be segmented before generating the spatiotemporal pattern. Although this may be achieved by eliminating the background using a temporal median filter, problems occur if the background moves even slightly.

The proposed system focuses on the use of periodic motion and skin colors to efficiently detect adult scenes. Adult scenes usually contain both large, periodically moving objects characterized by some skin color and backgrounds with contrasting colors that could be non-stationary and/or cluttered. The proposed approach directly extracts the spatiotemporal motion trajectories of objects (boundaries of thick 1D sine waves for periodic motion) from spatiotemporal patterns generated from videos. Thus, the proposed system does not require the segmentation of each frame into moving objects

[1] This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2012R1A1A2044892).

Soonhong Jung is with the School of Electrical Engineering, Korea University, 1, 5-ka, Anam-dong, Seongbuk-ku, Seoul 136-701, Korea. (e-mail: shjung@mpeg.korea.ac.kr).

Junsic Youn is with the School of Electrical Engineering, Korea University, 1, 5-ka, Anam-dong, Seongbuk-ku, Seoul 136-701, Korea. (e-mail: jsyoun@mpeg.korea.ac.kr).

Sanghoon Sull is with the School of Electrical Engineering, Korea University, 1, 5-ka, Anam-dong, Seongbuk-ku, Seoul 136-701, Korea. (e-mail: sull@korea.ac.kr).

Contributed Paper
Original manuscript received 09/22/14
Revised manuscript received 09/30/14
Current version published 01/09/15
Electronic version published 01/09/15.

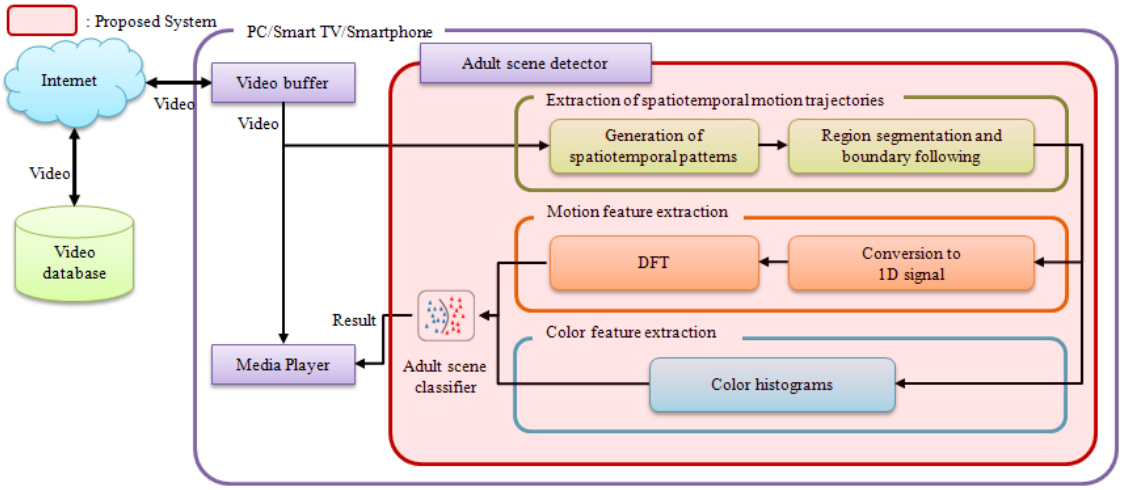

Fig. 1. The proposed system.

and their background. This is in contrast to existing methods which segment each frame based on color [9], motion [5]-[6] and temporal median filtering [11]-[13]. The proposed method allows for efficient motion segmentation and the robust detection of periodically moving skin-colored objects in front of non-stationary and/or cluttered backgrounds.

In this paper, spatiotemporal motion trajectories are extracted by first generating a spatiotemporal pattern from an input video, segmenting it into regions based on color, and then following the boundary of each region. By computing the vertical displacement of the spatiotemporal motion trajectory along the time axis of the spatiotemporal pattern, the spatiotemporal motion trajectory is then converted into a 1D signal which strongly resembles a sine wave when motion is periodic. To detect periodically moving skin-colored objects in a video, two types of feature vectors computed from i) the DFT coefficients of the 1D signal and ii) the colors in the region containing the spatiotemporal motion trajectory and its upper and lower neighbors are used as inputs to a classifier. Each step is computed using just the spatiotemporal pattern. Finally, the accuracy and efficiency of the proposed method are compared with those of six existing methods.

This paper is organized as follows. In Section II, the proposed algorithm is described. In Section III, the experimental results are presented. Finally, Section IV concludes the paper.

II. Proposed Algorithm

In this section, the details of the proposed system are described and summarized in Fig. 1.

A. Extraction of spatiotemporal motion trajectories

In this subsection, the steps for extracting spatiotemporal motion trajectories to locate the regions containing potential periodic motion in a spatiotemporal pattern are described.

1) Generation of spatiotemporal patterns

For periodic motion detection, a spatiotemporal pattern is first generated from a video [14]. The spatiotemporal pattern is obtained by accumulating a set of pixels sampled from each

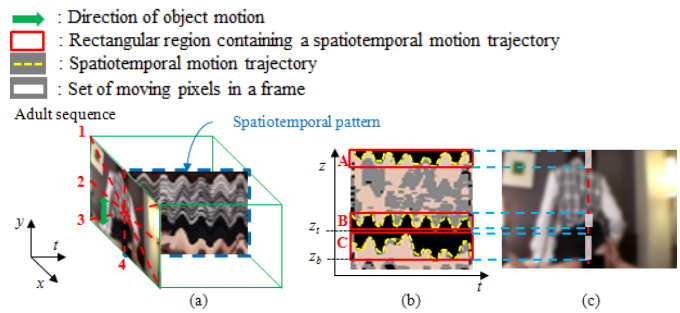

Fig. 2. Extraction of spatiotemporal motion trajectories for an adult sequence. (a) Four pixel sampling schemes and one of the corresponding spatiotemporal patterns. (b) Spatiotemporal motion trajectories extracted by color-based region segmentation and boundary following. (c) Sets of moving pixels where each set corresponds to a spatiotemporal motion trajectory.

frame along a time axis. Let $I(x, y, t)$ be the pixel value at location (x, y) and time t of a video. The spatiotemporal pattern $I_{ST}(z, t)$ is then defined as:

$$I_{ST}(z, t) = I(p(z), q(z), t), \tag{1}$$

where $p(z)=x$ and $q(z)=y$ are 1D functions of the independent variable z. Thus, the spatiotemporal pattern is a two-dimensional (2D) image consisting of pixels sampled at fixed positions in each frame of a video along a time axis.

The pixel sampling scheme for adult sequences covers four directions (horizontal, vertical, and the two diagonals), as shown in Fig. 2, because most large object motions can be effectively captured along at least one of these lines. For example, the dashed blue square in Fig. 2(a) indicates a spatiotemporal pattern constructed on scan line 4 along the time axis. Scene boundaries are detected by using the absolute values of the time differences in the spatiotemporal pattern. Each spatiotemporal pattern is divided into a set of non-overlapping unit video segments of duration T, which is experimentally set to 2.7 seconds. If the duration of a scene is shorter than T, the scene is ignored. Otherwise, each unit segment of the spatiotemporal pattern is analyzed to determine whether it contains periodic motion and skin color.

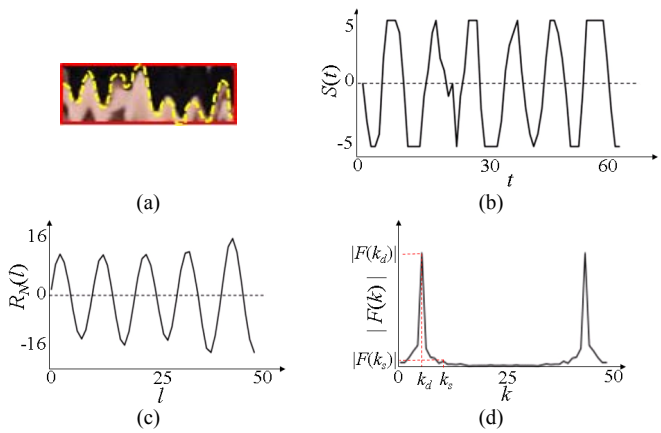

Fig. 3. Motion feature extraction. (a) The spatiotemporal motion trajectory in Region C of the adult sequence. (b) $S(t)$ obtained by computing the vertical displacement from Region C. (c) A 1D signal $R_N(l)$ of $S(t)$. (d) Magnitude of the DFT coefficients of $R_N(l)$.

2) Color-based region segmentation and boundary following

The spatiotemporal pattern in a given unit segment could exhibit a mixture of different periodic and/or non-periodic motion patterns appearing as thick waveforms propagating along the time axis as shown in Fig. 2(a). It can also be observed that the set of pixels along a sampling line passing through periodically moving objects corresponds to one or more thick 1D sine wave along the time axis. Thus, various motion characteristics such as (multiple) periodic or non-periodic motions can be distinguished once the boundaries of the thick waveforms (referred to as spatiotemporal motion trajectories) have been extracted. In this paper, the boundaries are extracted by simple color-based segmentation.

First, the spatiotemporal pattern or image is segmented using the k-means color clustering algorithm. The value of k is experimentally set to 3 where one of the initial values of the cluster centers is set to a skin color. This initial value is effective for the videos used in the experiments because, as shown in Fig. 2(a), skin color usually appears in a relatively large portion of at least one of the four spatiotemporal patterns, even if the frame itself contains only a small amount of skin color. Morphological algorithms such as closing and opening are then applied to the color-segmented regions to reduce noise. Since the spatiotemporal motion trajectories usually occur at the boundaries of differently colored regions, a simple boundary following algorithm [15] is applied to each of the segmented regions. A sample of the extracted spatiotemporal motion trajectories is shown in Fig. 2(b). Next, the smallest bounding box containing each sequence of connected boundary pixels is identified as a rectangular region. For example, the rectangular Region C in Fig. 2(b) is delineated by the top and bottom vertical coordinates z_t and z_b, respectively,

In the spatiotemporal pattern, pixels corresponding to stationary objects and the background appear as horizontal lines because the values of those pixels do not change along the time axis. In other words, a horizontal line corresponds to a single pixel at a constant frame coordinate that maintains a constant value across a succession of frames. Therefore, the rectangular regions of the spatiotemporal pattern corresponding to only a small amount of motion can be removed. In particular, the rectangular regions are removed from further consideration if their height is less than 10% of that of the entire spatiotemporal pattern or their width is less than half of the unit interval T. Fig. 2(c) shows sets of moving pixels associated with a spatiotemporal motion trajectory superimposed on a single frame.

The proposed approach to the segmentation of spatiotemporal patterns as shown in Fig. 2 alleviates the problem of distinguishing between moving foreground objects and a background which could be non-stationary or cluttered in a given frame.

B. Motion Feature Extraction

This subsection describes the steps necessary to extract motion features from a spatiotemporal motion trajectory which can then be used to detect periodic motion.

1) Conversion to 1D signal

A spatiotemporal motion trajectory extracted using the process above cannot generally be considered a function of time. Further, it does not produce a clean sine wave of periodic motion due to the imperfect region segmentation of the spatiotemporal pattern. Therefore, it is important to obtain a clean 1D signal function of time from the spatiotemporal motion trajectory. In this paper, the rectangular region in the spatiotemporal pattern is converted to the 1D intermediate signal $S(t)$ by computing the vertical displacement between $I_{ST}(\bullet, t)$ and $I_{ST}(\bullet, t+1)$ within the rectangular region containing the spatiotemporal motion trajectory as follows:

$$S(t) = \underset{k \in [-\frac{z_m}{3}, \frac{z_m}{3}]}{\operatorname{argmin}} \frac{1}{z_m - |k|} \left(\sum_{z=z_b}^{z_t-1} \left| I_{ST}(z,t) - I_{ST}(z-k,t+1) \right| \right), \tag{2}$$

where z_m $(=z_t - z_b)$ is the height of the rectangular region, and the integer k ranges between $\pm z_m/3$ because the boundary of interest is assumed to be smooth across consecutive frames.

To further refine $S(t)$, the autocorrelation of $S(t)$ is computed as follows:

$$R(l) = \frac{1}{T-1} \sum_{t=1}^{T-1} S(t) \cdot S(t+l), \tag{3}$$

where T is the length of $S(t)$, and l ranges from 1 to T - 1. Removing the DC value from $R(l)$ then obtains the normalized autocorrelation $R_N(l)$ corresponding to the desired 1D signal. In (3), the value of $R(l)$ increases as $S(t)$ and $S(t + l)$ become more correlated. Therefore, if a noisy $S(t)$ has periodicity, an improved waveform $R(l)$ can be obtained. $R(l)$ exhibits distinct peaks at multiples of the period $S(t)$. Because the values of $R_N(l)$ are not reliable at the start and finish of the rectangular region within the unit interval T, the values of $R_N(l)$ for both

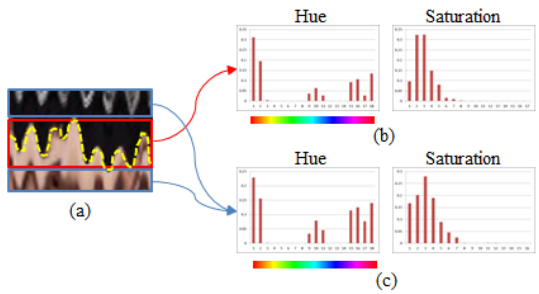

Fig. 4. Color feature extraction. (a) Region C containing the spatiotemporal motion trajectory of an adult sequence and its upper and lower neighboring regions indicated by blue lines. (b) Hue and saturation histograms of Region C. (c) Hue and saturation histograms of Region C's two neighboring regions.

the smallest and largest 12.5% of the values of l are not considered. Therefore, the length of $R_N(l)$ under consideration is equal to L, the size of which is equal to 75% of T - 1. Fig. 3(b) and (c) show the intermediate $S(t)$ and the 1D signal $R_N(l)$ for the spatiotemporal motion trajectory in Fig. 2(b), where it can be seen that the spatiotemporal motion trajectory is effectively converted to a 1D signal.

2) DFT of the 1D signal

To measure the periodicity, the 1D signal $R_N(l)$ is transformed to $F(k)$ by applying an L-point DFT. Denoting the magnitude of the dominant (first peak) frequency coefficient k_d and that of the second peak frequency coefficient k_s as $|F(k_d)|$ and $|F(k_s)|$, respectively, the degree of periodicity ρ is defined as follows:

$$\rho = 1 - \frac{|F(k_s)|}{|F(k_d)|}. \tag{4}$$

The values of ρ and k_d are input to the adult scene classifier to determine whether a periodic motion is present. Fig. 3(d) presents the magnitude of the DFT coefficients for the 1D signal $R_N(l)$.

C. Color Feature Extraction

Most existing methods [1]-[4], [7]-[10] for detecting adult scenes assume the presence of a large proportion of skin color in frames taken from adult scenes. Thus, these methods will have difficulty detecting adult scenes that contain only a small proportion of skin color. For typical adult video scenes, skin color appears in a relatively large portion of the spatiotemporal pattern generated along the sampling directions passing through the periodically moving body parts, even if they occupy only a small area of the entire frame, as shown in Fig. 2(a). Additionally, it is observed that in adult scenes, if the rectangular region contains skin color, at least one of its two neighboring regions is also likely to contain skin color, as displayed in Fig 4(a). Thus, the presence of skin color can be checked for the rectangular region containing the spatiotemporal motion trajectory and its two neighboring upper and lower rectangular regions, whose heights are empirically set to a half of the rectangular region as shown in

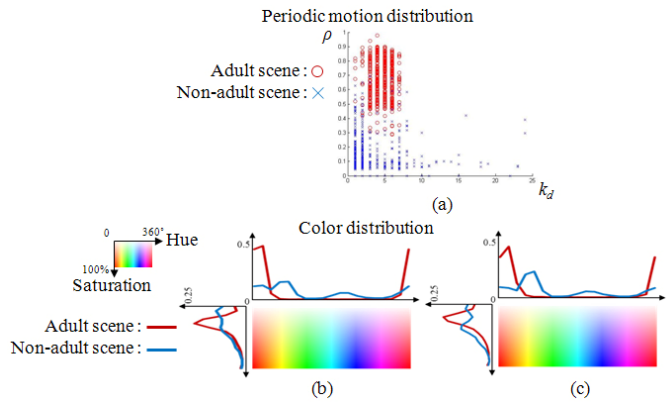

Fig. 5. Characteristics of adult scenes compared to non-adult scenes in the training set. (a) Distribution of motion features k_d and ρ. Distributions of averaged hue and saturation histograms for (b) the regions containing spatiotemporal motion trajectories and (c) the neighboring regions in the spatiotemporal pattern.

Fig. 4. Skin color is checked by constructing a 68-dimensional color feature vector consisting of two overlapping hue (18 bins) and saturation (16 bins) histograms computed from the rectangular region and its two neighboring regions, respectively. The color feature vector is also input to the adult scene classifier.

D. Adult Scene Classifier

In Fig. 5, the characteristics of adult scenes are compared with non-adult scenes in terms of the distribution of the motion and color features for a training set consisting of 1500 short video segments described in detail in Section III. It is observed that the use of simple thresholding will not work effectively, as shown in Fig. 5(a). It is also observed in Fig. 5(b) and (c) that the distributions of the averaged histograms for hue and saturation in adult scenes can be generally distinguished from those in non-adult scenes. The 2D motion feature vector (ρ, k_d) and a 68-dimensional color feature vector are used as inputs to the adult scene classifier based on an RBF-kernel support vector machine (SVM). If at least one of the four spatiotemporal patterns contains periodic motion and skin color, the video segment corresponding to the unit segment of the spatiotemporal pattern is classified as an adult scene.

III. Experimental Results

The performance of the proposed system was evaluated on a PC with a 2.8GHz CPU and 4GB of RAM. For comparison, six existing methods were also applied to the samples: gait motion detection [11], periodic motion detection [12]-[13], adult periodic motion detection using motion vectors [6], content-based obscene video recognition combining spatiotemporal and motion features based on skin color [9], and pornographic video content detection combining skin color, BOVW, MHIST, and MFCC [10]. For simplicity, minor modifications were applied to the existing methods. Instead of using the motion vectors encoded in a codec, motion vectors were independently computed for an 8×8 macroblock by employing the motion estimation method typically found in MPEG-2 or MPEG-4 encoders. The computation of motion

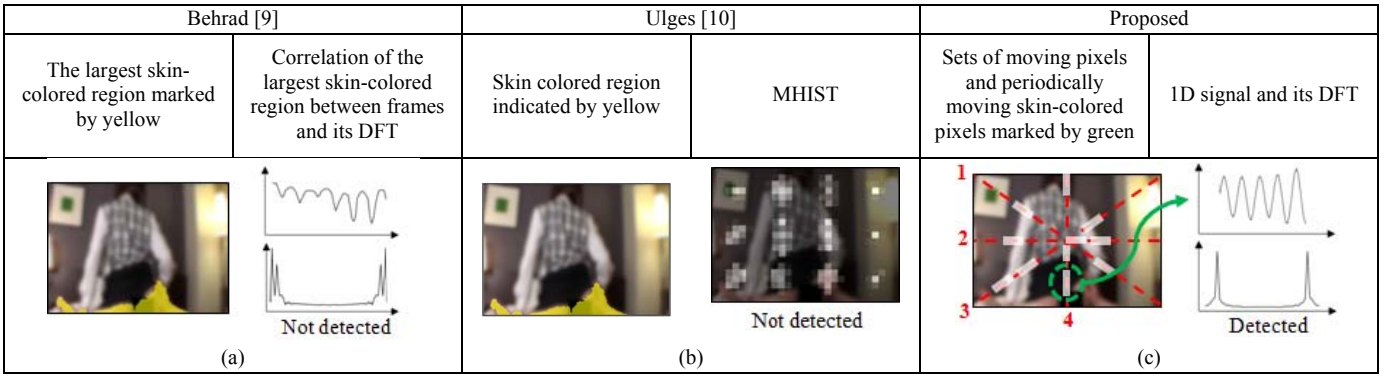

Fig. 6. Results of adult scene detection for an adult sequence using three methods (Behrad [9], Ulges [10], and the one proposed in this paper, respectively). (a)-(b) Not detected due to the small area of skin colors. (c) Detected. Each white region indicates a set of moving pixels corresponding to a spatiotemporal motion trajectory. The dashed green circle shows the detected sets of periodically moving skin-colored pixels.

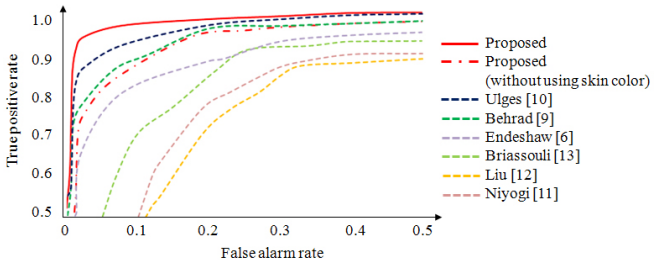

Fig. 7. ROC curve for detecting adult scenes.

vectors was not included when calculating the total processing time for the methods in [6] and [10]. Only visual cues in [10] were used, excluding the MFCC of an audio signal. To apply the periodic motion detection methods [11]-[13] for the purpose of detecting adult scenes, the values of ρ and k_d defined in the proposed method were used.

The training set consisted of a total of 1500 video segments with a duration of 2.7 seconds each (500 from adult videos and 1,000 from documentaries, sports programs, talk shows, comedies, and movies of various genres). The test was performed on a total of 70 hours of raw video collected from a variety of sources with a resolution of 640×480. A total of 18,313 scenes were extracted, out of which 1,103 scenes were detected to have contained at least one unit video segment containing an adult scene.

Fig. 6 illustrates the results using two recent methods [9]-[10] and the proposed method for a particular adult sequence. The methods using skin color [9]-[10] perform less effectively for adult scenes that contain a small proportion of skin color in each frame, as illustrated in Fig. 6(a) and (b). On the other hand, the proposed method provides satisfactory results in this situation as shown in Fig. 6(c). Note that in Fig. 6(c), although several sets of moving pixels are detected on the four directional pixel sampling schemes, only one set is selected in the end because it is the only one to contain both periodic motion and skin color. Fig. 7 shows the receiver operating characteristic (ROC) curve used in the detection of adult scenes for the six existing methods and the proposed method. It can be seen that the proposed method yields a better performance in terms of both true positive and false alarm rates. The proposed method also performs more effectively in the absence of skin color than the four existing methods that

Fig. 8. A false positive example and its spatiotemporal pattern.

do not use skin color as a component of classification [6], [11]-[13]. These four methods especially underperform when classifying scenes that contain non-stationary backgrounds. This is because Endeshaw's method [6] distinguishes the motion of objects in the foreground from objects in the background by using medians of motion vectors, and the other approaches [11]-[13] eliminate the background through temporal median filtering. It is experimentally observed that these four methods usually experience difficulty when the average magnitude of background motion is larger than 4 pixels. Behrad's method [9] uses the largest skin-colored region to estimate the periodicity, but errors occur when that area does not move periodically. Ulges' method [10] sometimes returns false positives for video scenes containing both skin color and dynamic motion, such as in mixed martial arts, because it does not estimate periodicity precisely. However, better results may be obtained if the MFCC of the accompanying audio signal is also used, as in their original approach.

Fig. 8 shows a video scene for which every method tested in the experiment, including the one proposed in this paper, returns a false positive. Video scenes similar to this could be more accurately detected by incorporating additional features, such as MFCCs.

Table I compares the processing times of the proposed method and the six existing methods [6], [9]-[13]. The methods based on spatiotemporal patterns [11]-[13], with the exception of the one proposed in this paper, require long computation times to eliminate the background using a temporal median filter because every pixel in each frame is involved in the computation. Briassouli's method [13] in particular takes longer due to the computation required to find the optimal μ. Endeshaw's method [6] is more efficient than the proposed one because it uses the motion vectors from a compressed video stream. However, it yields lower true

TABLE I
COMPARISON OF EFFICIENCY IN DETECTING ADULT SCENES

Methods	Processing time for 6 hours video
Niyogi [11]	2 hours 32 minutes
Liu [12]	2 hours 26 minutes
Briassouli [13]	5 hours 33 minutes
Endeshaw [6]	1 minute 21 seconds
Behrad [9]	18 minute 20 seconds
Ulges [10]	3 hours 17 minutes
Proposed (without using skin color)	**9 minutes 9 seconds**
Proposed	**9 minutes 30 seconds**

positive and higher false alarm rates, especially for video scenes with non-stationary backgrounds due to its problem with segmentation. Behrad's method [9] considers every frame when extracting regions containing skin color to construct the spatiotemporal volumes of skin-colored regions. Ulges' method [10] has higher computational complexity because it calculates the DCT coefficients for every local image patch in key frames. On the other hand, the method proposed in this paper requires little time to analyze skin color. These experiments show that the proposed approach is efficient and yields improved true positive and false alarm rates.

IV. CONCLUSION

In this paper, an efficient real-time system has been proposed to detect adult scenes based on spatiotemporal patterns. The spatiotemporal motion trajectories are efficiently extracted by the simple color-based region segmentation of spatiotemporal patterns, even in the presence of a non-stationary and/or cluttered background. The feature vectors from the DFT of the 1D signals converted from each spatiotemporal motion trajectory and the presence of skin color are used as inputs to a classifier. Because each step is efficiently and reliably processed based only on spatiotemporal patterns, the experimental results demonstrate that the proposed method exhibits superior performance. This performance could be improved further by the addition of features such as audio cues.

REFERENCES

[1] M. J. Jones, and J. M. Rehg, "Statistical color models with application to skin detection," *International Journal of Computer Vision*. vol. 46, no. 1, pp. 81-96, 2002.

[2] D. A. Forsyth, and M. Fleck, "Automatic detection of human nudes," *International Journal of Computer Vision*. vol. 32, no. 1, pp. 63-77. Aug. 1999.

[3] T. Deselaers, L. Pimenidis, H. Ney, "Bag-of-visual-words models for adult image classification and filtering," *Proceedings of the 19th International Conference on Pattern Recognition, ICPR 2008*, pp. 1-4, 2008.

[4] A. P. B. Lopes, S. E. F. de Avila, A. N. A. Peixoto, R. S. Oliveira, A. de Araújo, "A bag-of-features approach based on hue-sift descriptor for nude detection," *Proceedings of the 17th European Signal Processing Conference*. *EUSIPCO 2009*, pp. 1552-1556, 2009.

[5] N. Rea, G. Lacey, C. Lambe, and R. Dahyot, "Multimodal periodicity analysis for illicit content detection in videos," *3rd European Conference on Visual Media Production*. pp. 106-114, 2006.

[6] T. Endeshaw, J. Garcia and A. Jakobsson, "Classification of indecent videos by low complexity repetitive motion detection," *Applied Imagery Pattern Recognition Workshop*, pp. 1-7, 2008.

[7] S. Lee, W. Shim, and S. Kim, "Hierarchical system for objectionable video detection," *IEEE Trans. Consumer Electron.*, vol. 55, no. 2, pp. 677-684, May. 2009.

[8] C. Kim, O. Kwon, and S. Choi, "A practical system for detecting obscene videos," *IEEE Trans. Consumer Electron.*, vol. 57, no. 2, pp. 646-650, May. 2011.

[9] A. Behrad, M. Salehpour, M. Chaderian, M. Saiedi, and N. Barati, "Content-based obscene video recognition by combining 3D spatiotemporal and motion-based features," *EURASIP Journal on Image and Video Processing*. Dec. 2012.

[10] A. Ulges, C. Schulze, D. Borth, and A. Stahl, "Pornography detection in video benefits (a lot) from a multi-modal approach," *ACM Multimedia*, pp. 21-26, 2012.

[11] S. A. Niyogi, and E. H. Adelson, "Analyzing gait with spatiotemporal surfaces," *IEEE Workshop on Motion of Non-Rigid and Articulated Objects*, 1994.

[12] F. Liu, and R. W. Picard, "Finding periodicity in space and time," *International Conference on Computer Vision*, pp. 376-383, Jan. 1998.

[13] A. Briassouli and N. Ahuja, "Extraction and analysis of multiple periodic motions in video sequences," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 7, pp. 1244-1261, Jul. 2007.

[14] H. Kim, J. Lee, J. Yang, S. Sull, and S. M. Song, "Visual rhythm and shot verification," *Multimedia Tools and Applications*, pp. 227-245, 2001.

[15] R. C. Gonzalez and R. E. Woods, "Digital image processing," Prentice Hall, 2002.

BIOGRAPHIES

Soonhong Jung received the B.S. degree in Electronic Engineering from Korea University, Seoul, Korea in 2007, and the M.S. degree in Electrical Engineering from Korea University, Seoul, Korea in 2009. He is currently working toward the Ph.D. degree in Electrical Engineering at Korea University. His current research interests are image/video signal processing, video coding, and N-screen image/video service.

Junsic Youn received the B.S. degree in Electronic Engineering from Korea University, Seoul, Korea in 2009. He is currently working toward the M.S. and Ph.D. joint degree in Electrical Engineering at Korea University. His current research interests are image processing and computer vision.

Sanghoon Sull (S'79-M'81-SM'13) received the B.S. degree with honors in electronics engineering from the Seoul National University, Korea, in 1981, the M.S. degree in electrical engineering from the Korea Advanced Institute of Science and Technology in 1983, and Ph.D. degree in electrical and computer engineering from the University of Illinois, Urbana-Champaign, in 1993. In 1983-1986, he was with the Korea Broadcasting Systems, working on the development of the teletext system. In 1994-1996, he conducted research on motion analysis at the NASA Ames Research Center. In 1996-1997, he conducted research on video indexing/browsing and was involved in the development of the IBM DB2 Video Extender at the IBM Almaden Research Center. He joined the School of Electrical Engineering at the Korea University as an Assistant Professor in 1997 and is currently a Professor. His current research interests include image/video processing, computer vision, data analysis, and their applications to 3D TV and smart phone. Dr. Sull is currently an associate editor for IEEE Transactions on Circuits and Systems for Video Technology.