Technical Disclosure Commons

Defensive Publications Series

June 03, 2016

PORNOGRAPHY DETECTION IN VIDEO USING CHARACTERISTIC MOTION PATTERNS

Rahul Sukthankar

Shumeet Baluja

PORNOGRAPHY DETECTION IN VIDEO USING CHARACTERISTIC MOTION PATTERNS

Given the large volume of video uploads to various Internet websites and services, there is increasing demand for automated methods for automatic identification of inappropriate content, such as pornography. Traditional approaches attack the problem primarily from an image or "appearance-based" perspective: e.g., based on the quantity of flesh-colored pixels not associated with faces. Such methods can generate a high number of false positives since they may incorrectly flag videos (e.g., medical documentaries, scenes on a beach, even skin-colored objects (furniture, fabric, etc.) as pornographic.

We present a mechanism for detecting pornographic content in a video based on an analysis of characteristic motions in the video. This motion-based mechanism is complimentary to appearance-based approaches, and therefore can be employed in combination with such approaches, or alternatively could be employed independent of appearance-based approaches. In addition, the motion-based mechanism is complimentary to audio-based techniques (e.g., classifying sounds or background music as being correlated with pornographic content, etc.), and therefore can be employed in combination with such techniques, or alternatively could be employed independent of audio-based techniques.

In one implementation, our motion-based mechanism employs a first phase comprising unsupervised learning, and then a second phase comprising supervised learning. In the first phase, a collection of pornographic videos is temporally segmented into intervals, and unsupervised learning discovers characteristic motion patterns and clusters the intervals by characteristic motion type. In one implementation, characteristic motion patterns are discovered in the first phase by analyzing the relative displacements of large numbers of ordered trajectory pairs over time. In accordance with this technique, the first trajectory of a given pair defines a reference frame for measuring the motion of the second trajectory of the pair, and each trajectory is associated with a particular moving portion of an object.

In the second phase, a supervised-learning classifier (e.g., a support vector machine [SVM], a neural network, a decision tree, etc.) is trained to learn which of the clusters identified in the first phase are highly correlated with pornographic content. The trained classifier can then be used to predict whether content in a video outside of the training set (e.g., a newly-uploaded video, etc.) contains pornographic content.

It should be noted that the trained classifier can be used on its own to detect pornographic content, or can be employed in conjunction with one or more other techniques, such as appearance-based approaches based on skin color, object recognition, etc., audio-based techniques based on sounds and music, etc., and so forth. It should further be noted that in some implementations, the second phase might be performed by a human rather than a machine-based classifier. More particularly, once the clusters have been identified by the unsupervised learning of the first phase, the clusters can be quickly visualized and tagged by a human, as opposed to laboriously tagging videos from scratch without the benefit of automated clustering.

Our motion-based mechanism has several advantages over previous techniques. First, our motion-based mechanism does not require videos to be temporally segmented, and therefore can be applied to full-length videos containing multiple topics. Accordingly, shorter segments of pornography within a longer video can be detected, thereby enabling detection of pornographic content that is inserted (often maliciously) in between portions of benign content.

Second, in contrast to approaches that seek to match video content to one or more human models, our motion-based mechanism does not rely on any assumptions concerning the subjects (e.g., actors, etc.) in the video. Consequently, the mechanism will work regardless of the number of subjects, and is robust with respect to partial occlusion.

Third, our motion-based mechanism is robust with respect to variations in appearance (e.g., color manipulation, black and white content, etc.). It is thus capable of detecting characteristic motion patterns in cartoons (e.g., adult anime, adult cartoons, etc.), video games, and videos with false or altered color, in contrast to appearance-based approaches based on skin tone and color.

Figure 1 depicts a flow diagram of a method for detecting pornographic content in videos, in accordance with some implementations. The method begins by identifying a collection of videos having pornographic content (also referred to simply as "pornographic videos," for convenience). The method then executes a first preprocessing phase in which characteristic motion patterns are discovered and clustering is performed. An implementation of a method for performing characteristic motion pattern discovery and clustering is described in detail below with respect to Figure 2.

Next, a second preprocessing phase is performed, in which a classifier is trained to associate the clusters identified in Phase I with pornographic content. A newly-uploaded video V is then identified, and motion patterns in video V are determined (e.g., via trajectory analysis, etc.). A motion pattern descriptor (e.g., a feature vector, etc.) is then provided as input to the trained classifier to obtain a pornography score for video V. Optionally, the motion-based pornography score is then combined with one or more scores obtained by other techniques (e.g., appearance-based, audio-based, etc.). At this point, the combined score (or simply the motion-based score) can be compared to a threshold in order to predict whether video V contains pornographic content.

Figure 2 depicts a flow diagram of a method for characteristic motion pattern discovery and clustering, in accordance with some implementations. First, the collection of pornographic videos is temporally segmented into intervals. Next, characteristic motion patterns are discovered via unsupervised learning. An implementation of a method for discovering characteristic motion patterns via unsupervised learning is described in detail below with respect to Figure 3. Next, the intervals obtained in the first step of the method are clustered by characteristic motion type.

Figure 3 depicts a flow diagram of a method for discovering characteristic motion patterns via unsupervised learning, in accordance with some implementations. First, trajectories for moving portions of objects are determined. Next, trajectory pairs $P_1...P_N$ are established, where pair $P_j = (R_j , T_j)$ and where trajectory R_j is a reference frame for trajectory T_j. Next, characteristic motion patterns are obtained via an analysis of relative displacements of trajectory pairs $P_1...P_N$ over time. This technique is described in detail in L. Del Pero, S. Ricco, R. Sukthankar, V. Ferrari, "Articulated motion discovery using pairs of trajectories", Proceedings of the International Conference on Computer Vision and Pattern Recognition, 2015.

ABSTRACT

A mechanism for detecting pornographic content in a video based on an analysis of characteristic motions in the video. The motion-based mechanism is complimentary to appearance-based and audio-based approaches, and therefore can be employed in combination with such approaches, or by itself. In one implementation, the mechanism employs a first phase of unsupervised learning, and a second phase of supervised learning. In the first phase, characteristic motion patterns are discovered and clustering is performed. In one implementation, characteristic motion patterns are discovered by analyzing the relative displacements of large numbers of ordered trajectory pairs over time. In the second phase, a classifier is trained to associate the clusters identified in the first phase with pornographic content. A motion pattern descriptor (e.g., a feature vector, etc.) for a video (e.g., a newly-uploaded video, etc.) is obtained and is provided to the trained classifier to obtain a pornography score for the video .

Keywords: video, pornography, motion, machine learning, trajectory pairs

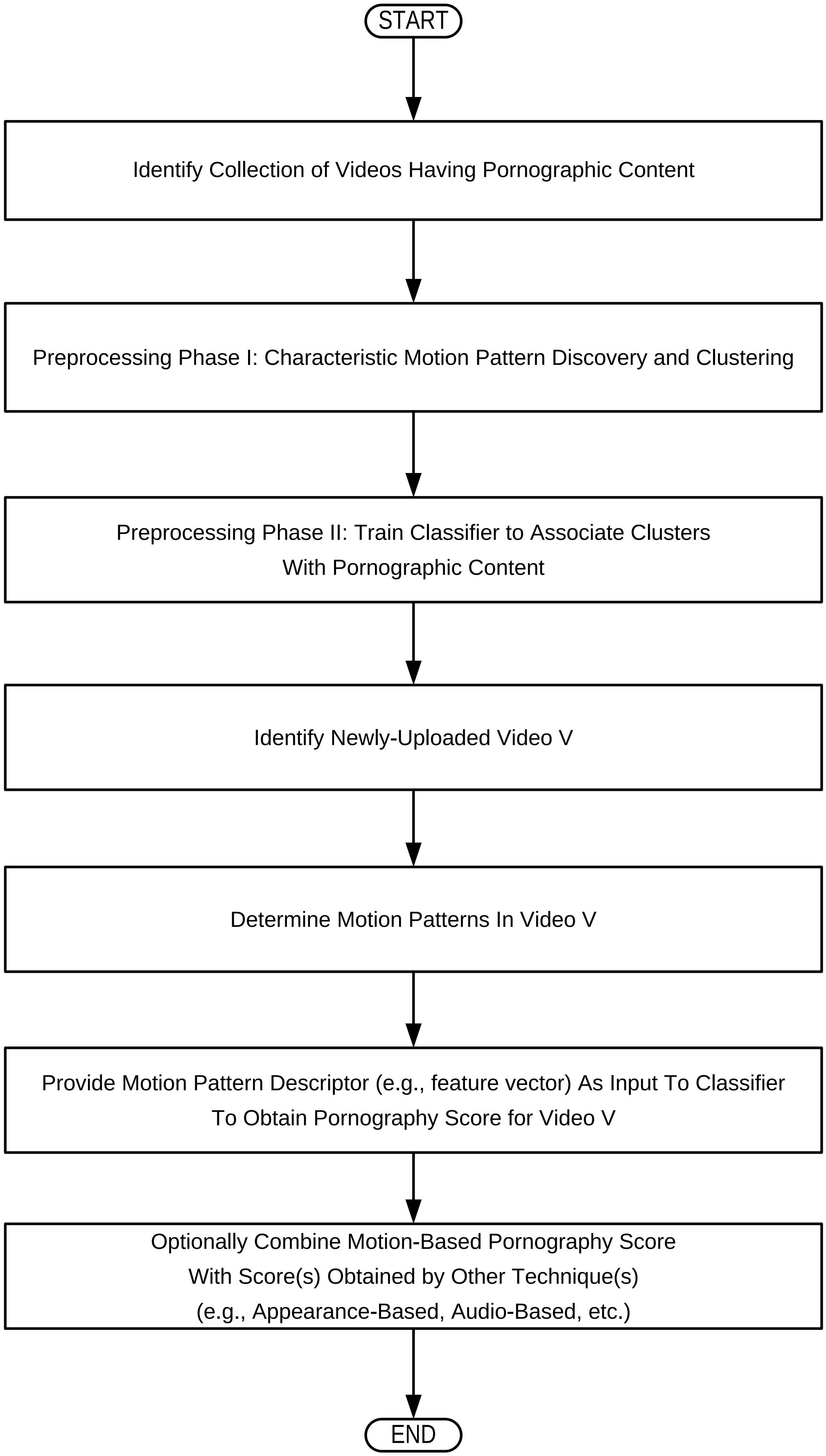
START
Identify Collection of Videos Having Pornographic Content
Preprocessing Phase I: Characteristic Motion Pattern Discovery and Clustering
Preprocessing Phase II: Train Classifier to Associate Clusters With Pornographic Content
Identify Newly-Uploaded Video V
Determine Motion Patterns In Video V
Provide Motion Pattern Descriptor (e.g., feature vector) As Input To Classifier To Obtain Pornography Score for Video V
Optionally Combine Motion-Based Pornography Score With Score(s) Obtained by Other Technique(s) (e.g., Appearance-Based, Audio-Based, etc.)
END

FIG. 1

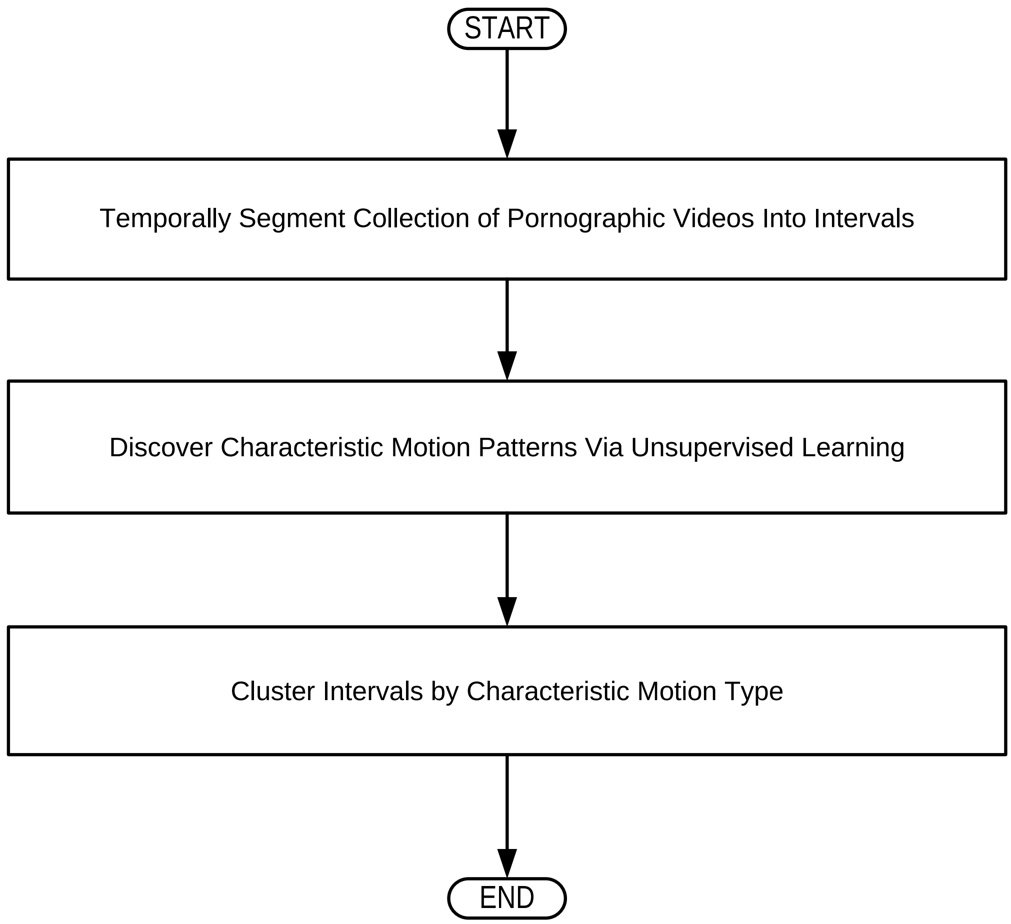
START
Temporally Segment Collection of Pornographic Videos Into Intervals
Discover Characteristic Motion Patterns Via Unsupervised Learning
Cluster Intervals by Characteristic Motion Type
END

FIG. 2

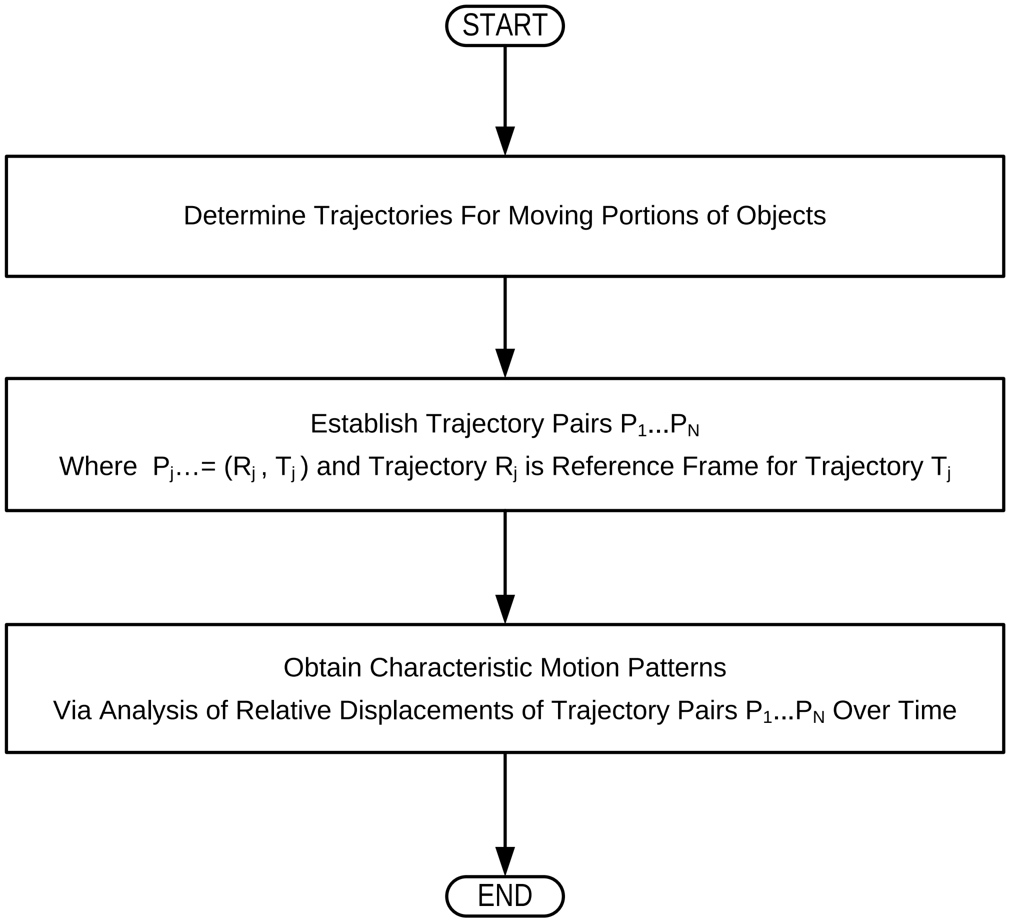
START
Determine Trajectories For Moving Portions of Objects
Establish Trajectory Pairs $P_1 ... P_N$
Where $P_j ... = (R_j, T_j)$ and Trajectory R_j is Reference Frame for Trajectory T_j
Obtain Characteristic Motion Patterns
Via Analysis of Relative Displacements of Trajectory Pairs $P_1 ... P_N$ Over Time
END

FIG. 3