# Fast Objectionable Videos Detection Algorithm Based on Motion Vector Field

Lv Guiqing
School of Information Science & Engineering
Lanzhou University line
Lanzhou ,China
Lvguiqing2010@163.com

Qu Zhiyi
School of Information Science & Engineering
Lanzhou University line
Lanzhou ,China
quzy@lzu.edu.cn

***Abstract*—Aim at large amount and complex features of network video data, this paper proposed a fast algorithm of detecting pornographic video in streaming media. Combining the information of motion vector field and the features of pornographic behavior trajectory, we could judge that weather a video is pornographic or not without decoding it. The experiment show that the algorithm recognition rate is reached 80%, and is robust for network video detection environment**

***Keywords-Motion Vector; Feature Extraction; Reciprocating Motion; Histogram; H.264***

## I. INTRODUCTION

With the growing popularity of the network, it has entered ordinary people's homes and become part of our lives. Because of the positive and the nagetive of the network information, how to create a healthy and green network environment, especially to protect the healthy growth of young people and strengthen the network information monitor with pornographic information filtering, attracted people's attention. Pornography pictures detection has been relatively mature technology at home and abroad, which is detected and interceped using face recognition, gesture recognition, skin detection and texture matching technology; As for pornography videos detection, there are still many problems to be solved, the current generic method uses the lens of video segmentation, key frame extraction, and skin texture detection and setting the threshold to achieve a certain effect, but a huge amount of calculation of such detection, for the massive network video information is difficult to carry out a rapid and timely screening. To solve these problems, this paper proposes the use of the motion vector field information, starting from the characteristics of the human trajectory, rapid detection of suspicious pornography video without completely decoded.

## II. ALGORITHM

This article propose a method for extracting motion vector field from the H.264 compression video streams without uncompressing the compressed video data directly, and then denoising and accumulating the motion vector field to obtain reliable temporal-spatial motion vector field, at last we determine whether there are a large number of simple reciprocating trajectory in human body after seeking motion histogram and using the the HMM training models. The block diagram of the algorithm is shown in Figure 1.

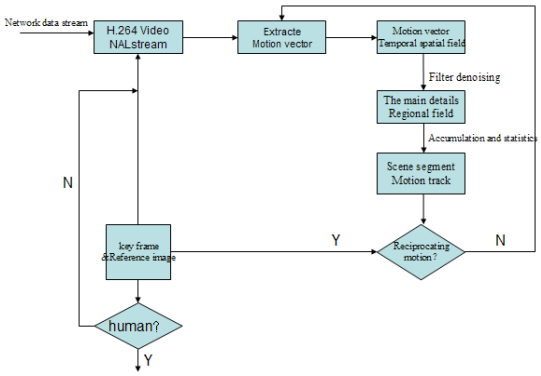



Figure 1 the block diagram of the algorithm

## III. SPECIFIC ANALYSIS

To detemine a video is pornography or not involving two main problems: First, the detection of human body, this involves human gesture recognition (feature extraction and module-based model identification); Second, the determination. of what kind of motion(human behavior recognition), which involves the identification of motion process(module-based or probability and statistics-based). Based on characteristics of H.264 video codecs and the possibility of rapid detection of a pornography video under the network environmentm, we preliminaryly propose the method of screening whether there exists reciprocating motion using the histogram paragraph of motion vector, after extracting keyframe on the video segment where exists reciprocating motion and determining whether there exists the human body, we can basically determine whether this video section contains pornography content

H.264 video codec standard draft doesn't make clear each module specific implementations, just making demands on syntactic and coming up with a standard on every module port stream. H.264 video codec standard proposed VCL layer and the NAL layer, the NAL layer is mainly to solve the problem of network transmission, dividing into two relatively independent transmission channels: the reference image set channel and slice group channel. It also comes up with five levels of each independent but interrelated in the syntax, of which slice layer becomes quite independent syntactic coding, each frame consists of one or more different slice, each slice contains one or more microblock and each microblock can be divided into different subblocks according to the degree

of richness of the contents. Macroblock predictive coding contains two forms: intra-frame prediction coding and inter-frame prediction coding. According to our needs, we need to extract the relative displacement of the interframe macroblock, and then judge whether the macroblock or motion region exists reciprocating in the time domain.

### A. *Principle Analysis*

In order to compress the data stream as far as possible, H.264 coding choose the reference picture as close as possible to (match) the current picture to reduce residuals, and it generally choose 1-5 reference pictures to achieve the security of the network transmission, this provided favorable conditions for us to detect. We can assume that the reference picture set is the keyframes set or lens field set, in which repeated motion of objects in the field, or the positive and negative expression of the motion vector constitutes a rich content of the judgment contains combining with the keyframe picture.

### B. *Macroblock (Subblock) Motion Vector Arithmetic*

As independent syntactic layer, there are five different types of slice in H.264 codec, while each type of the slice containing the corresponding type of macroblock and subblock, and each macroblock presented different movement tendency relative to some reference time at different time. In order to extract the motion vector effectively and get a true reflection of the object motion trajectory, macroblock (subblock) motion vector should be converted into a unified three-dimensional coordinate system to calculate. I macroblock and SI macroblock are coded with the method of intra-slice prediction, their motion vector arithmetic should be carried out in a two-dimensional space domain rather then in a three-dimensional time-space domain.As for B macroblock, P macroblock and SP macroblock, their motion vector arithmetic should be carried out in a unified coordinate system based on a different reference picture sequence. Such as time domain frame arrangement 'IBBPBBPBBIPPB ...', motion vector accumulation, in principle, should be accumulate anchored to one I frame. As shown:

$$\begin{cases} Mv_1=Mv_0+d_1(x_1,y_1) \\ Mv_2=Mv_0+d_2(x_2,y_2) \\ Mv_3=Mv_0+d_3(x_3,y_3) \\ \cdots\ \cdots \\ Mv_n=Mv_0+d_n(x_n,y_n) \end{cases} \cdots\cdots\rightarrow \overrightarrow{MV}=\{\vec{d}_1,\vec{d}_2-\vec{d}_1,\vec{d}_3-\vec{d}_2,\ldots\ \vec{d}_n-\vec{d}_{n-1}\}$$

1、a macroblock MV value based on the same reference frame

$$\begin{cases} Mv_1=Mv_0+d_1(x_1,y_1) \\ Mv_2=Mv_1+d_2(x_2,y_2) \\ Mv_3=Mv_2+d_3(x_3,y_3) \\ \cdots\ \cdots \\ Mv_n=Mv_{n-1}+d_n(x_n,y_n) \end{cases} \cdots\cdots\rightarrow \overrightarrow{MV}=\{\vec{d}_1,\vec{d}_2,\vec{d}_3,\ldots\ \vec{d}_n\}$$

From i frame, the motion direction of some region start to reverse.

$$\overrightarrow{MV}=\{\vec{d}_1,\vec{d}_2,\ldots,\vec{d}_i\ \ldots,\vec{d}_n\}$$

2、a macro block MV expression based on the macro block of the previous frame

$$\begin{cases} Mv_1=Mv_a+d_1(x_1,y_1) \\ Mv_2=Mv_a+d_2(x_2,y_2) \\ \cdots\ \cdots \\ Mv_i=Mv_a+d_i(x_i,y_i) \\ Mv_j=Mv_b+dj(x_j,y_j) \\ \cdots\ \cdots \\ Mv_n=Mv_b+d_n(x_n,y_n) \end{cases}$$

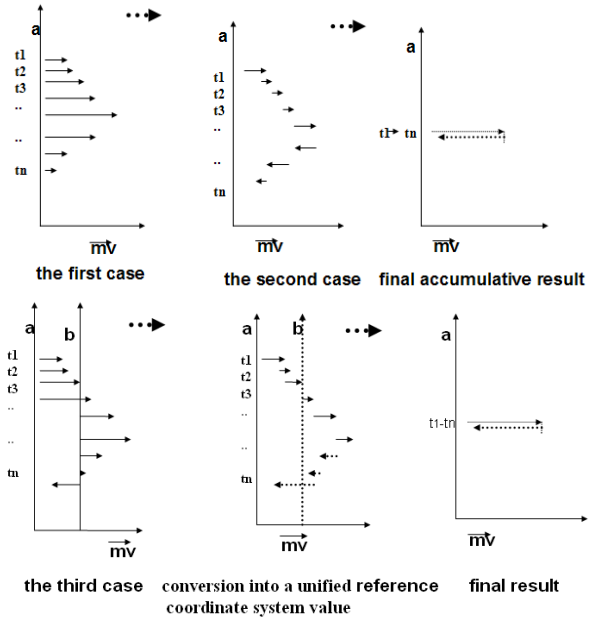
$$\overrightarrow{MV}=\{\vec{d}1,\vec{d}2-\vec{d}1,\ldots,\vec{d}i-\vec{d}i\text{-}1,\vec{b}-\vec{a}(\vec{d}j,\ldots,\vec{d}n-\vec{d}n\text{-}1)\}$$

3、a macroblock MV expression based on the macroblocks of two different reference frames

In the experiment, the extracted motion vector field is a 3D data：

$$\prod_{Mv}=\begin{Bmatrix} mv_{11}\ mv_{12}\ mv_{13}\ldots mv_{1n} \\ mv_{21}\ mv_{22}\ mv_{23}\ldots mv_{2n} \\ \cdots \\ \cdots \\ mv_{m1}\ mv_{m2}\ mv_{m3}\ldots mv_{mn} \end{Bmatrix} \begin{Bmatrix} d_1(x_i,y_j) \\ d_2(x_i,y_j) \\ \cdots \\ \cdots \\ d_t(x_i,y_j) \end{Bmatrix} (i,j)\in(m,n)$$

Ideally, if the extracted motion vectors is exactly the entire cycle of the reciprocating movement, in the above cases, a macroblock motion vector accumulation as shown above figures：



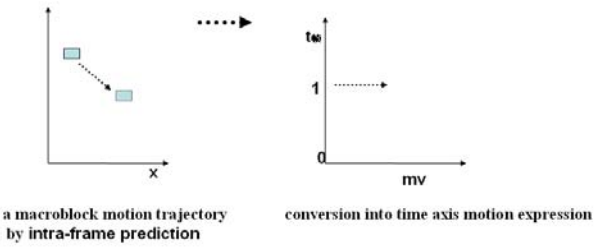



Figure 2 a macroblock motion vector accumulation

Thus can be found, if a macro block exists reciprocating movement, the vector value extracted by simple accumulation does not reflect the true motion trajectory, only the extracted vector converted into the unified coordinate system value can effectively express motion trajectory. When counting and accumulating the motion vector value in a unified coordinate system, we can continue to accumulate the motion vector value when the movement trend of the current macroblock MV and the previous frame macroblock MV is in the same direction region, when it is not in the same direction region, we should add variables to store their motion vector value, at last we can obtain the actual object running track in whole time domain field referencing to these variables value and location.As for the I macro block, we should divide it into two cases, one case is it is a reference macroblock basis itself, the MV is set to 0; the other is the reference macroblock is the adjacent macroblock of the slice, MV value is considered to be equal in spatial domain and both of its dimension and value increases by 1.

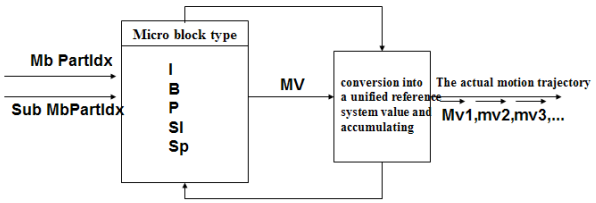



Figure 3 Motion vector extraction and conversion process

The following chart (the main macro block MV of 25 frames statistual results):

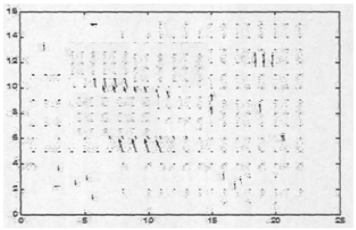


Figure 4 The original data accumulation

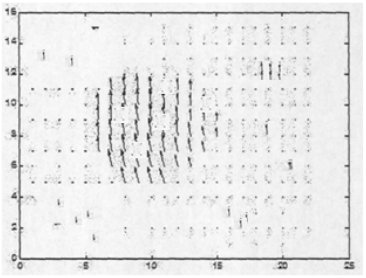


Figure 5 The result after the data conversion

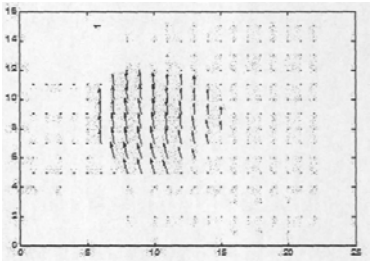


Figure 6 the result After denoising

As can be seen, H.264 motion vector expression seemed sparse due to the adoption of a macroblock coding. That is conducive to the compressed video stream. However, as for vector detection, it increased difficulty. In order to enhance our detection effectiveness, we also need ilrlage denoising and accumulation in time-space domain. After the effective accumulation, the motion vector strength is improved, and after high-pass filtering ilrlage denoising, we obtained the major region where exists motion vector (i.e. detail describes macroblock). According to the spatial domain, we can see the movement trajectory and trend of the main macroblock and the subblock, thus we can preliminary judge its characteristics.

## IV. MOTION CHARACTERISTICS AND ORIENTATION HISTOGRAM

From the motion vector field information we can see that the dense degree of the motion vector field reflects the strength of motion vector and expresses the macroblock details part, and the direction of the motion vector field reflects object motion trend. We can get direction histogram conveniently using the two-dimension numerical statistics of the motion vector. The range of the orientation is in 0~2 π, each 45 degrees of which is divided into an interval, a total of eight intervals, thereby forming a eight numerical interval statistics. As the interval (1, 5; 2, 6; 3, 7; 4, 8) is repeated a certain number, we can conclude that the movement characteristics is reciprocating.

According to human body motion frequency feature in a pornographic video (expected early period reciprocating 1cycle every 1 second, climax period about reciprocating 2 cycle every 1 second). Assert that video frequency is 30fps, every 15-30 frames as a cycle, such as a video of 6 minutes has 180 frames, an average of 30 frame has 1 - 2 times of the reciprocating motion, It is judged to be a pornographic video

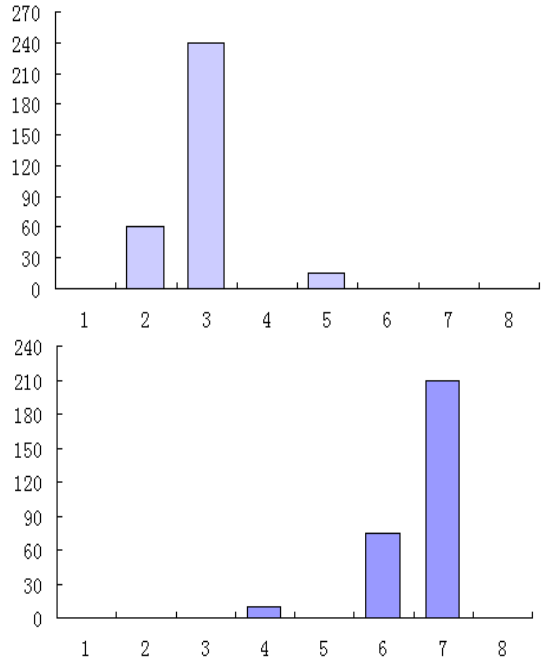



Figure 7 the motion vector statistic histogram of the macroblock &subblock in one second

## V. NAKED HUMAN BODY IDENTIFICATION

In the experiment, we combine face recognition with skin color detection method to judge a pornographic video. When judging that a given video exists reciprocating motion, we truncate a frame image of this section randomly, and perform pretreatment, after we detect human face, we can judge whether the proportion of skin color of the connected region achieves the degree of pornography.

## VI. SUMMARY

In the experiment, we count motion vector value from different motion vector macroblock reference system, unifying reference system of every video segment solving the problem of cumulative error of motion vector value, thus providing more accurate reference for later judgment. Based on the motion trajectory and skin color detection, can achieve a satisfactory detection result to a certain extent.

## REFERENCES

[1] BiHoujie.A new generation of video compression coding standard:H.264/AVC[M].Beijing:Posts & Telecom Press,2005 (in Chinese)

[2] ISO MPEG-4 Part10,H.264/MPEG-4AVC[S]

[3] Shi Pengju,Sun Guangmin,LiYan.The DSP implementation of moving object detecting and tracking based on MPEG-4[C] IEEE Industrial Electronics.Piscataway,NJ:IEEE,2006:3279-3283.

[4] Zhao X H. Research 0n Content Based Erotic Image Filtering Technique[D]. Jilin : Jilin University, 2003. .

[5] Chen Juan ,Xia Jun,Yin Hanchun Extracting Motion Vectors in Compressed Video Sequence,Chinese Journal of electron Devices Dec. 2006 V01. 29 No. 4 1342-1345

[6] XIE Zhiyang, SHI Ping The Research and Implementation of Obj ectionable Internet Video Information Filtering System, JOURNAL OF COMMUNICATION UNIVERSITY OF CHINA(SCIENCE AND TECHNOLOGY) Dec, 2009 67-71

[7] DingWenrui, YangHua,Moving object detecting technique based on motion vectorin H.264 coding, Journal of Beijing University of Aeron autics and Astron autics, February 2010,202-205

[8] Liu Yanmin,Research on Porn Video Detecting Algorithom Based On Motion Features, Ms D Thesis:LanZhou University,2010