ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

Content Based Multimedia Copy Protection System for the Cloud

Vaishali Dewar, Priya Pise

M. E Student, Department of Computer, Indira College of Engineering and Management, Savitribai Phule Pune University, Pune, Maharashtra, India

Professor, Department of Computer, Indira College of Engineering and Management, Savitribai Phule Pune University, Pune, Maharashtra, India

ABSTRACT: These days creating multimedia and multimedia editing has become very easy dueto easily available processing tools. More over hosting websites are freely available.So it is obvious that we encounter situations wherein many videos and images thatare copyrighted are getting duplicated easily. Such duplication is not only illegalbut also it causes a tremendous loss to original content holder. If at all we try todetect such duplicated content, process gets very complex because of large sizes andlarge quantity of multimedia files present on web. So this process becomes complex and computationally expensive.Content based multimedia copy detection providesa robust and optimized mechanism which can identify matching percentage amonghuge multimedia copies from the content itself without having any dependency onany software. In the technique of content based copy detection, signatures are createdfrom content itself before uploading on cloud like Picasa. In the proposed systemwhenever new Multimedia object is uploaded on the cloud firstly its signature iscreated and if the signature matches with stored database signature then this willtrap the site which duplicates copyrighted copy without rights.

KEYWORDS: Content based copy detection, cryptosteganography, multimedia objects,SIFT.

I. INTRODUCTION

Rapid growth in the multimedia technologies has made very easy to keep and access large amount of multimedia files in storage networks. There are lots of easy editing and publishing mechanism available which makes duplication of multimedia data feature very easy which can cause the violation of digital rights. So, copy rights security becomes a critical problem for the multimedia data over the cloud. All these hurdles and challenges has created the necessity of developing the new mechanism for detection of the copied video over the internet. Duplication of copyrighted materials makes huge loss to content owners. Consider an example where a party creates a video and sales its copyrights to a hosting party say YouTube which pays for it per view of the video on YouTube to content owner. But if that copyrighted material is leaked and hosted on some different hosting site then it will make loss to Content owner as well as YouTube. So this copyrights stealing need to be caught.

Many problem such as storage management and copyright violations arises because of the reason that people more commonly upload numerous videos for the sake of Business promotion or community sharing.

I) Storage management issue

If the identical copies of videos are kept in the storage system then the process becomes very costly as it requires large space for storing it. Moreover, the process of retrieving that multiple copies of videos becomes more time consuming. Therefore, if the duplicate copies of videos are identified with using an easy mechanism, then surely the Storage management issues can be fixed effectively.

II) Copyright violations issue -

It has been very easy now a days for creating transformed videos and then upload them on internet. This has immensely affected the multimedia group or broadcasting agencies with great deal of loss. Manually for a human operator, it seems very tiresome and almost impossible to detect the presence of duplicated version of original video content.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

All these hurdles and challenges has created the necessity of developing the new mechanism for detection of the copied videos.Therefore, this mechanism of detecting video copy has come out as an outstanding way to minimize major piracy and copyright issues. Due to the large size of the available multimedia content over the Internet and the complexity of comparing content to identify copies, detection of these pirated content on Internet has turned out to be very complex and computationally expensive operation
Recent studies for identifying copyright piracy issues are mainly classified into –

i. Watermarking based and
ii. Content based copy detection.

These techniques has its own advantages and disadvantages. Watermarking inserts useful information i.e. copyright information and also it has low complexity in order to access this information from multimedia files. Using this technique some distortion like Fourier transformation is added in multimedia copy and to play or view the multimedia object in original form this distortion need to be removed by some software or players, which makes dependent on those players. So it cannot be helpful for online videos which are being uploaded in hundreds and thousands in number per day.

Whereas Content Based Copy Detection (CBCD) algorithms does not depend on any information insertion completely and are resistant to edition and fusion. These CBCD algorithms identifies differentiating characteristics from the media copy itself and based on these feature detects the duplication so CBCD mechanism is more fault proof and more effective approach for piracy detection on internet. Provided that an enough information is available in video to create its unique differentiating factor; this means video itself helps to identify itself.

We propose a new design, in this along with CBCD mechanism we will use cryptosteganography mechanism to hide secrete copyright information in multimedia copy to fasten the copyright detection mechanism.

II. RELATED WORK

Nowadays creating multimedia and performing operations i.e. editing, processing tools has become very easy. More over hosting websites are freely available. So it is obvious that we encounter situations wherein many videos, images, and music clips that are copyrighted are getting duplicated easily. Such duplication is not only illegal but also it causes a tremendous loss to original content holder. If at all we try to detect such duplicated content, process gets very complex because of large sizes and large quantity of multimedia files present on web. So this process becomes complex and computationally expensive. M. Hefeeda in 2015 presented a new technique of content based copy protection system with the help of multi-cloud which can support different multimedia files like audio, video etc.

Many studies mainly focused on identifying varieties of visual features which can be used for video copy detection systems and came up with a summary showing latest advancements in the process of copyrights piracy detections.

Previously a Video sequence matching method based on graphs and a segmentation was specifically used, because of better stability and its ability to discriminate its local features [3]. But graph based matching is computationally expensive.

Comparison with existing system:

The issues and complexity in protecting the multimedia objects of various kinds has grabbed the interest of many researchers and research institutes. Watermarking is one of these approaches to solve this problem. In this mechanism a copyrights identifying content is embedded in the object of multimedia itself and then a method is used which searches for this content so that we can determine whether the multimedia object is the authenticated one or not. Watermarking needs not only embedding multimedia objects in watermarks but also needs a mechanisms/systems which can check those and specify whether correct watermarks are present in them or not. This technique is not susceptible for any distortion in the file.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

Disadvantages:

1. Watermarking requires embedding watermarks in the multimedia objects and requires to search for secrete information in files and verify the presence of correct secrete information added using watermarking technique in them.
2. It requires specific software to play the multimedia, these software removes the watermarks from images.

III. PROPOSED ALGORITHM

Watermarking technique was used for the same purpose of piracy detection. Using this technique some distortion like Fourier transformation is added in multimedia copy and to play or view the multimedia object in original form this distortion need to be removed by some softwares or players, which makes dependent on those players.
We propose a new design that is based on using the cloud of multiple software resources which collectively works together to divide the computing task. The main concept which is being used is "Content-based copy detection" (CBCD) mechanism which is an effective technique for detecting multimedia copyrights piracy on cloud. Different types of multimedia like audio, video or images can be protected using the CBCD mechanism.

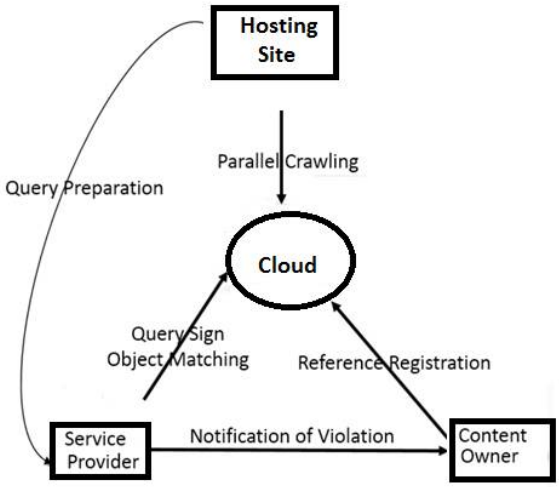


Fig. 1. System Architecture

CBCD – Content Based Copy Detection

Content based copy detection (CBCD) technique is independent of any software based dependency and identifies copyrighted materials from its content itself. In this technique signatures are created from content itself using SIFT feature. Multimedia objects which owners want to protect are mentioned by them. Then, system creates signatures from that file and stores them in the database. The Crawl component once a day downloads recently uploaded files from the hosting sites (like YouTube). Then signatures are also created from these query resulted objects. Signature of query result objects are then matched with all database signatures. If signature matches then this will trap the site which duplicates copyrighted copy without rights.

This is effective mechanism, but to fasten this process we will use Audio video cryptosteganography which adds secrete information in multimedia object itself which can neither be removed nor can be identified visually as watermarks are visible. This secrete information is considered as copyrighted information before forming the digital signature. If this secrete information matches with secrete information from database then all contents in the multimedia will be cross verified using CBCD and in this way copyrighted material will be identified.

Content owners like Yash Raj Film industry, hosting sites like YouTube, or service providers like Disney any of these three parties can install this system to detect copyright piracy on internet.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

1. Distributed Database: It stores signatures of multimedia files which content owners want to protect.
2. Copyrighted Signature Registration: Audio video cryptosteganography add secrete information in multimedia object itself along with inserting the authorized signature in database.
3. Query Preparation: Here crawler downloads relevant multimedia objects in response to search query. Then signatures is created from objects downloaded from online sites, which are called query signatures. It then uploads these signatures for object matching component;
4. Signature comparison: Here query signatures and secrete information versus reference signatures and secrete information in the distributed index are compared to find potential match. It also notify content owners if piracy is detected. In this process firstly secrete information will be extracted and if it is available then key frame matching steps can be avoided and in this way piracy checking can be made fast.

Proposed system mainly focuses on video/images piracy identification. Even though these multimedia objects are edited then proposed algorithm identifies piracy detection based on percentage of matched content of edited multimedia objects with respect original multimedia objects. If contents are matched then Service Provider will come to know about the content duplication.

Mechanism For Video Content Duplication Detection:

Proposed algorithm works on identifying image content duplication detection. Same algorithm can be used for identifying video content duplication because video is sequences of images i.e frames. However video may have redundant frames which we can ignore while content matching. From the video key-frames will be extracted. Logic behind extracting key-frame is that only consider scene changing frames. That means from video only those frames will be extracted which are identifying frames in the video. These array of key-frames will be passed to image Content Duplication Detection algorithm.

Mechanism For Image Content Duplication Detection:

Proposed algorithm works on identifying image content duplication detection. Every image has some identifying features in it which makes it visually appealing. These identifying features ares used by human brains to compare two images. Same concept will be used for image comparison in the proposed algorithm. These identifying features are called key-features. In the proposed method SIFT key-features are extracted and the signature is calculated from the RGB content and byte value of the SIFT points. Proposed system uses SIFT features as it is more robust to image compression.

Signature Creation:

The proposed system is designed to handle two types of multimedia objects - video and image. The system abstracts the details of different media objects into signatures. The signature creation and comparison component is media specific, while other parts of the system do not depend on the media type. SIFT features are extracted from image.

Scale Invariant Feature Transform(SIFT):SIFT employs Difference of Gaussian to detect local maxima values and these interest points are described by gradient histogram based on their orientations. Proposed system uses SIFT descriptor due to its good stability and discriminating ability. SIFT feature performs well among local feature category and is robust to scale variation, rotation, noise, affine transformations.

Hiding Secrete Information Using Steganography:

Steganography is the method of hiding any secret information in multimedia objects like audio, video and images. It aims to hide secrete information behind the original cover file. Add signature's half bits information as an secret information using steganography mechanism in file itself at the end of file, i.e at last to lenghth of files. If any of the multimedia object identifies registered stego-signature as is then whole process of image/video content duplication process can be skipped and thus time will saved.

Advantages of Proposed System:

1. Supports Large-scale Datasets
2. Cost Efficiency

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

IV. ALGORITHM

KeyFrame Extraction Algorithm:

The algorithm compare and calculate the similarity of each video frames to consider whether there is a change in the scenery or not. If there is a change, we break the video here and finally we will break the video into shots.

```
ForEach frame  i in video
  binValueOfFrame-i = frame-i
 binValueOfFrame-j = frame-i+1
   if binValueOfFrame-i – binValueOfFrame-j > Threshold
  Add frame i in keyFramesList
   else
     keyFramesList.add(i+1)
End
```

Signature Creation Algorithm

Step 1:

Divide video frames as left and right parts frames and evaluate Visual Descriptors for each of them.Each frame is considered as image. Around each pixel descriptor i is computed in the image, which has a location of (xi, yi). (SIFT features) using formula –

DiL= (fil, fi2,, …fiF), i = 1, 2, ….Ln,

DiR= (fjl, fj2,, …fjF), j = 1, 2, ….Rn,

Step 2:

Divide each left and right frames into Blocks of same number - (NxM)

Step 3:

Compute the Visual Descriptors deviation for each descriptor in the left frame with respect to right frame and identify the nearest descriptor as -

$$DiL - DjR = \sqrt{(fi1 - fj1)2 + \ldots + (fiF - fjF)2}$$

Step 4:

Compute Block Deviation – calculate the block deviation of each part of NxM subframes in the left frame with respect to subframe in the right frame. Say Sbi, where i is block index and calculated as.

$$\sqrt{((xi - xj) / Wb)2 + ((yi - yj) / Hb)2}$$

Step 5: Create Signature. The signature from these corresponding frames is:
(Sb1,Sb2, . . . , SbN).

V. MATHEMATICAL MODEL

Mathematical Model:

$$m = \sum_{i=1}^{M} \sum_{j=1}^{N} Cb_{ij} / MN$$

Where,

1. m is matching percentage normalized w.r.t. Image size

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

2. Cbij is byte value Weight per block i.e SIFT point in Image
3. M is Height
4. N is Width

Set Theory:

A] Let S be the system ,
S→ {……..}
Identify I as the input
I→{m}
Where ,
m→ input multimedia object for signature creation

B] Identify p1, p1 as a processes ,
P1→{k,is,f,l,s,si }
Where,
k→ key frames
is→ image segmentation
f→ feature extraction
l→ local feature vector is generated
s→signature creation
si→ signature indexing

P2→{q,qs,qsm}
Where,
q→querying for related multimedia
qs→queried objects signature creation
qsm→queried signature matching
S = {I,P1,P2}

C] Identify O as Output
O = {j}
Where ,
j→queried object signature matched with stored indexed signature.

S = {I,P1,P2,O}

D] Identify A as case of success
A = {x}
Where ,
x→relevant multimedia object given as input
S = {I,P,O,A}

E] Identify Y as case of failure
Y = {o}
Where o → irrelevant multimedia object given as input
S = {I,P,O,A,Y}

VI. SIMULATION RESULTS

- network. For each frame of the query video, the signature is computed and the closest signatures to it are retrieved from the distributed index.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

For example, if frame x in the query video matches frame y in the reference video, we expect frame x+1 in the query video to match frame y+1 in the reference video.

- We measure the performance in terms of two basic metrics:
  precision (percentage of returned copies that are true copies)
  recall (percentage of returned copies that are false copies).
- So we can check whether a match declared by the system is a true match or not by viewing the copy personally.
- In the system as we go on increasing threshold value, percentage of true detection goes on increasing i.e. precision value increased whereas percentage of false detection is decreased.
- With below input iumages, when uploaded, system identifies 76% matching.

- Result given by the matching algorithm is

Dataset consist of video and images of 8 different editing/transformation categories which itself include 200 to 250 samples. Each images and video are transformed with different mutimedia transformation mechanisms which are listed below.

The performance of the whole system is evaluated with a large dataset of almost 1GB multimedia files. Dataset used in the evaluation process is from web-crawled misc database used in WBIIS for CBIR Techniques. These multimedia objects are from different categories, and have diverse sizes, durations, resolutions, and frame rates, file formats etc.

1. Image cutting
2. Image Resize
3. File Format Change
4. File Compression
5. Watermarked images/ Watermarked video
6. Video Frame dropping
7. Row-interleaved
8. Column-interleaved

For each above categories, we evaluated precision and recall by varying the threshold value. Below are the precision and recall average value among all categories:

1. Precision (percentage of returned copies that are true copies)

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

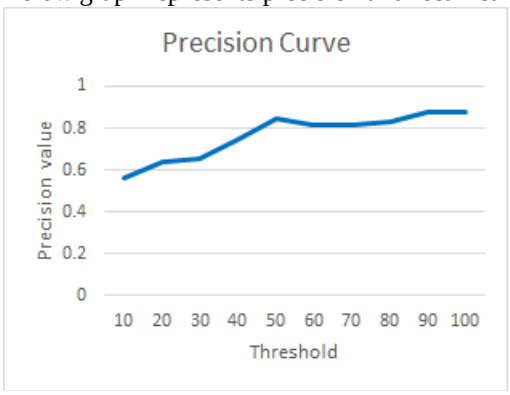
Vol. 4, Issue 10, October 2016

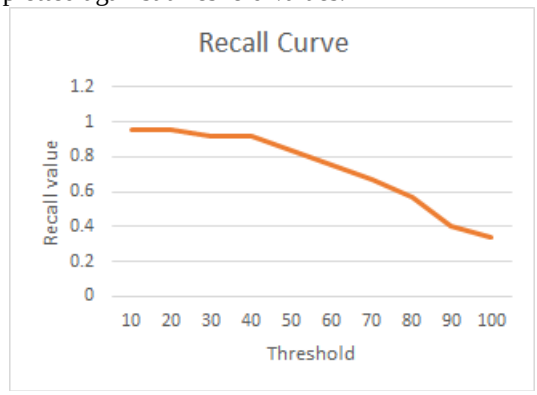
2. Recall (percentage of returned copies that are false copies).

Threshold	Recall	Precision
10	0.958333	0.5625
20	0.958333	0.6375
30	0.916667	0.654167
40	0.916667	0.745833
50	0.833333	0.85
60	0.75	0.8125
70	0.666667	0.8125
80	0.572917	0.833333
90	0.40625	0.875
100	0.34375	0.875

Fig: Result Table

Below graph represents precision and recall curve plotted against threshold values.





To get this PR curve,the threshold is changed from 10 to 100, and computed the precision and recall for each threshold. PR curves are a standard evaluation method in image retrieval, as they contain rich information and can easily be read. The results clearly show that system can achieve both high precision and recall. For example, a precision of 100% with a recall of more than 80% can be achieved. The results clearly describes that how the precision and recall vary with the threshold parameter s. The results show that method can achieve precision and recall values of more than 80% for a wide range of multimedia editing categories. This means that system does not only provide high accuracy, but it is very sensitive to the threshold s, so service provider can change the threshild as per the sevirity of the issue which is an internal system parameter. In other words, the system administrator can accurately fine tune s depending the system needs.

VII. CONCLUSION AND FUTURE WORK

Content based multimedia copy protection system is a novel design for video/image copyright protection system on cloud. The system is based on inserting digital signatures using image cryptosteganography which adds secrete information in multimedia object itself as a copyright information in the media itself and authorizing it so that it can be used to fasten the content based copy detection method. This mechanism will help to detect copyrighted multimedia stealing and will help to stop loss to content owner due to illegally duplicating copyrighted material. However this

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 10, October 2016

work is not covering the piracy checking in the areas where live moments are telecast on online telecasting sites. This area needs be extended further.

REFERENCES

1. Mohamed Hefeeda, Tarek ElGamal, Kiana Calagari and Ahmed Abdelsadek "Cloud-based Multimedia Content Protection System". IEEE, 1520-9210, 2015
2. Sunil K Moon, "Application of data hiding in Audio- Video using Anti Forensic Technique for Authentication and Data Security", IEEE, 978-1-4799-2572-8/14/$31.00 c, 2014.
3. A.PerumalRaja, B.Venkadesan, R.Rajakumar "Efficient Framework for Video Copy Detection Using Segmentation and Graph-Based Video Sequence Matching". IEEE, 2250-3153, Volume 4,9, September 2014.
4. Hong Liu, Hong Lu, and XiangyangXue,"A Segmentation and Graph-Based Video Sequence Matching Method for Video Copy Detection". IEEE, 2250-1706, Volume 25, NO. 8, AUGUST 2013.
5. DhanalakshmiSrinivasan,"An Effective Approach for Video Copy Detection and Identification of Misbehaving Users". 0975-9646, Volume. 4 (6) ,863-867, 2013.
6. Vishwa Gupta, Parisa Darvish, Langis Gagnon, Gilles Boulianne,"CONTENT-BASED VIDEO COPY DETECTION USING NEAREST-NEIGHBOR MAPPING". 4673-0382, Vol.1 (6) ,2013.
7. Shikui Wei, Yao Zhao, Ce Zhu, ChangshengXu,andZhenfengZhu,"Frame Fusion for Video Copy Detection". IEEE, 305-732, VOL. 21, NO. 1, JANUARY 2011
8. Hye-Jeong Cho, Yeo-Song Lee, Chae-Bong Sohn, Kwang-Sue Chung, and Seoung-Jun Oh,"A NOVEL VIDEO COPY DETECTION METHOD BASED ON STATISTICAL ANALYSIS". VOL. 1, NO. 978-1-4244-4291, 2009
9. Mani MalekEsmaeili, MehrdadFatourechi, and Rabab KreidiehWard,"A Robust and Fast Video Copy Detection System Using Content-Based Finger printing". VOL. 6, NO. 1, MARCH 2011
10. Vishwa Gupta, Parisa Darvish, Langis Gagnon, Gilles Boulianne,"VIDEO COPY DETECTION USING INCLINED VIDEO TOMOGRAPHY AND BAG-OF-VISUAL-WORDS". 305-732, Vol.1 (6) ,2012.