ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

*INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING*

*National Conference on Advanced Innovation in Engineering and Technology (NCAIET-2015)*

*Alva's Institute of Engineering and Technology, Moodbidri*

*Vol. 3, Special Issue 1, April 2015*

# Indoor-Outdoor Classification of Audio-Video and Detection of Tampered Audio

**Shijitha P U[1], Nagaraja Hebbar N[2]**

2nd Year M.Tech, Srinivas Institute of Technology, India[1]

Associate Professor, Srinivas Institute of Technology, India[2]

**Abstract:** An improved method for indoor-outdoor classification of audio and video is presented. In second part it explains the detection of tampered audio. For indoor-outdoor classification of audio RT60 approach is used. This method calculates the reverberation time and decay curve of impulse response of the given audio track. The reverberation time value is compared with the threshold value for performing the classification. The contrast of hue component of HSV image of video frame is calculated for video classification. The tampering detection is performed by comparing the reverberation time of original and corrupted audio. The proposed solution proves to be useful in identifying fake audio tracks. It also helps in scene classification in videos ie to detect whether the given video is indoor or outdoor.

**Keywords:** Include RT60, HSV.

## I. INTRODUCTION

Audio tampering detection is quite challenging and active research area. Due to the presence of editing tools the origin and content in video can be altered. Many editing tools are present using which audio and video can be edited. Some case it is useful but if misused it can result in serious problem. Video and audio are considered as evidence in law suits but due to the presence of editing tools it can be edited easily, due to this reason it cannot be considered as trustworthy evidence. The audio and video tool available [3] is used for many purposes. It includes editing, playing, recording, converting audio and video. In movies audio and video editing has many advantages. It is useful in dubbing, adding canned laugh, editing video.

The video tampering can be done by different ways. The video frame in video file can be tampered by replacing the video with another video sequences or by replacing the video with still images that repeats in time [6]. Sometime video sequence from some other video is also used for tampering. However altering a video is relatively easy task due to the development of many user friendly tools. This software helps people in forging a video easily, the forging is done with high perfection which make it difficult for detecting. Similarly audio is tampered either by replacing some part of audio with the audio from some other audio. The audio can also be tampered by adding background noise or by replacing a word in the audio by some other word.

The one technique used for audio tampering detection is audio finger printing [7]. Another approach is audio hashing technique [8].

The proposed method calculates reverberation time RT60 [2, 5] for indoor-outdoor classification and for tampering detection. T60 measures reverberation time and integrated impulse decay curve of impulse response. If the reverberation time value is less, the audio is indoor else the audio is outdoor. For tampering detection two audio are checked and their reverberation time value and decay curve is compared. If both are same for two audio it is concluded that no tampering has occurred else it is tampered. The existence of sound, because of multiple reflections from different surfaces in a room causes distortion in recorded sound. This distortion is referred as reverberation time. The difference in RT60 value can be used for comparison. Each room will be having different reverberation time value.

For indoor/outdoor classification of video hue value of HSV image is compared [4]. If hue value is greater than the threshold value the scene is classified as indoor or else it is an outdoor video. The RGB model used does not consider the luminance effect, which is one of the main drawbacks. To overcome this drawback HSV model is used for indoor-outdoor classification of the given videos.

## II. PROPOSED METHOD

Two different methods are proposed here. The first explains indoor-outdoor classification of audio and video. And the second explains the detection of tampered audio.

### *A.* Indoor-outdoor classification.

The acquired video is converted to video frames. These frames are quantized to avoid the redundant information. The frames are then converted to HSV (Hue-Saturation-value) images. Here hue refers to depth, Saturation refers to purity and value refers to intensity of the colors. This

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

*INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING*

*National Conference on Advanced Innovation in Engineering and Technology (NCAIET-2015)*

*Alva's Institute of Engineering and Technology, Moodbidri*

*Vol. 3, Special Issue 1, April 2015*

model is used in many applications for the identification of different objects.

From the HSV image hue component and v components are separated. After that H component is quantized into 8 levels. Quantization is a lossy compression which is achieved by compressing a range of values to a single quantum value. Color quantization reduces the number of colors used in an image thus it helps in reducing the file size. The value of hue is compared if the hue value is greater than 0.25 the videos are classified as indoor else it is an outdoor video.

The block diagram reported in Fig 1 explains how it works. The user input the video and the videos are converted to video frames. From the entire frame 10 frames are selected and are converted to HSV images. The H and S components are separated and H values compared and if it is above the given threshold video is classified as indoor .If the hue value is below the given threshold it is an outdoor video.

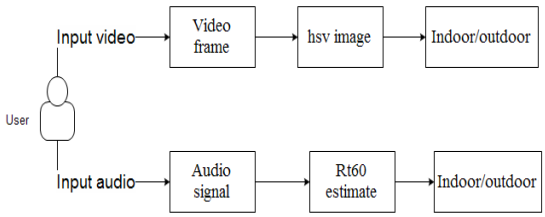


Fig 1: block diagram for indoor-outdoor

Audio are stored as audio signals. The frequency of audio signals is in the rage of human hearing frequency ie 20 to 20,000Hz. The audio classification is done by estimating RT60 value. It measures the reverberation time. The reverberation is due to multiple reflections of audio signals. The existence of sound, due to multiple reflections from different surfaces in a room causes disturbance in recorded sound. This distortion is referred as reverberation time. If the reverberation time value is less it is an indoor audio otherwise it is an outdoor audio. The decay is created due to the sound absorbed by different object in the space-including furniture, wall, air, people, table etc. It depends on the absorption of sound by the different object as well as the shape of the room.

The block diagram given above Fig: 1 describes the audio analysis. The user will input the audio track. The audio signal is plotted for given audio. Followed by this RT60 estimate is calculated to obtain the reverberation time. The reverberation time value is calculated ad its value is compared with the given threshold. If the reverberation time value is less it is indoor else it is outdoor. The RT60 estimate returns reverberation time and the decay curve of impulse response.

*B.* Tampering detection.

This section describes the method of detecting the tampered audio. The reverberation time and decay curve is calculated separated for both audio. The input audio is selected .The RT60 value is calculated and is displayed for both audio separately. The impulse decay curve graph of impulse response for both audio is drawn. After that comparison of reverberation time and graph is done. If the values are same for both audio it is not tampered.

If the audio is tampered there will be difference in the values. On comparing the output and verifying the result the message is displayed whether it is a tampered audio or not.

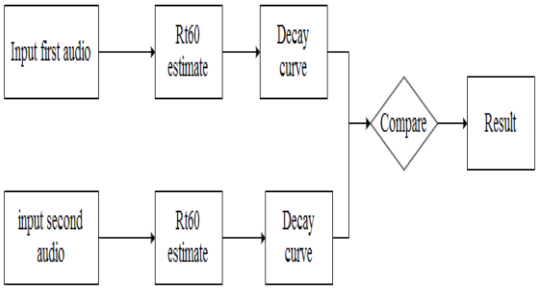


Fig 2: block diagram for tampering detection

The Fig 2 is for the method explained above. Here the first audio and the second audio are given as input. The reverberation time ad RT60 value is calculated ad graph is drawn for both. The compare module will compare both the output and finally the result is given

## III. EXPERIMENTAL RESULT

The proposed detector successfully classifies the indoor and outdoor videos. The first set includes 10 videos recorded in known environment (indoor & outdoor) .The table1 report the result of video analysis. The hue value of each videos and estimated environment based on hue value is shown.

TABLE I. INDOOR-OUTDOOR CLASSIFICATION OF VIDEO

| VIDEO NAME | REAL ENVIRONMENT | HUE VALUE | ESTIMATED ENVIRONMENT |
|---|---|---|---|
| V1 | INDOOR | .814 | INDOOR |
| V2 | OUTDOOR | .244 | OUTDOOR |
| V3 | INDOOR | .414 | INDOOR |
| V4 | INDOOR | .444 | INDOOR |
| V5 | OUTDOOR | .090 | OUTDOOR |
| V6 | INDOOR | .596 | INDOOR |
| V7 | OUTDOOR | .244 | OUTDOOR |
| V8 | OUTDOOR | .041 | OUTDOOR |

Next table demonstrate the indoor-outdoor classification of audio. The reverberation time for the audio and the decay curve obtained are shown.

ISSN (Online) 2321 – 2004
ISSN (Print) 2321 – 5526

*INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH IN ELECTRICAL, ELECTRONICS, INSTRUMENTATION AND CONTROL ENGINEERING*

*National Conference on Advanced Innovation in Engineering and Technology (NCAIET-2015)*

*Alva's Institute of Engineering and Technology, Moodbidri*

*Vol. 3, Special Issue 1, April 2015*

The test is done on several audio for classifying. The decay curve of impulse is shown in the Fig: 3.

TABLE I. INDOOR-OUTDOOR CLASSIFICATIO N OF AUDIO

| VIDEO NAME | REAL ENVIRONMENT | RT60 VALUE | ESTIMATED ENVIRONMENT |
|---|---|---|---|
| A1 | INDOOR | 10312 | INDOOR |
| A2 | INDOOR | 6133 | INDOOR |
| A3 | OUTDOOR | 26499 | OUTDOOR |
| A4 | INDOOR | 13832 | INDOOR |
| A5 | OUTDOOR | 13571 | OUTDOOR |
| A6 | OUTDOOR | 87920 | OUTDOOR |
| A7 | INDOOR | 6963 | INDOOR |

The Fig: 3 is the decay curve graph of the impulse response

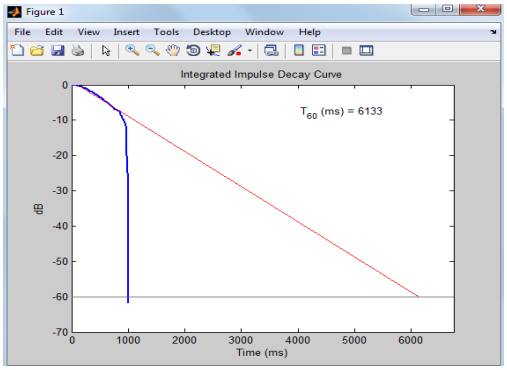


Fig: 3: Impulse decay curve

In the graph the red lie is the reverberation time and blue is the decay curve.

The result of tampering detection is given in table III.

TABLE III. TAMPERING DETECTION

| RT60 OF FIRST AUDIO | RT60 OF SECOND AUDIO | TAMPERED |
|---|---|---|
| 6963 | 6800 | YES |
| 6133 | 6133 | NO |
| 10312 | 10312 | NO |
| 6145 | 6100 | YES |
| 10312 | 10300 | YES |

## IV. CONCLUSIONS

The indoor-outdoor classification of audio-video and detection of tampered audio is done. The experimental result shows that the detector is capable of performing both the analysis.

## ACKNOWLEDGMENT

I would like to express my gratitude to my guide, The Head of the Department and The principal of srinivas institute of technology for the help and support given by them for completing this work.

## REFERENCES

[1]. Simone Milani, Pier Francesco Piazza, Paolo Bestagini, Stefano Tubaro, "Audio tampering detection using multimodal feature", Dipartimento di Elettronica, Informazione eBioingegneria, Politecnico di Milano, Milano, Italy.

[2]. H. Malik and H. Farid, "Audio forensics from acoustic reverberation, "in Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP 2010), March 2010, pp. 1710–1713.

[3]. S. Milani, M. Fontani, P. Bestagini, M. Barni, A.Piva, M. Tagliasacchi, and S. Tubaro, "An overview on video forensic,"APSIPA Transactions on Signal and Information Processing, vol. 1, pp. 1–18, 2012.

[4]. punith kumar m b, dr. p s puttaswamy, "scene classification based on feature extraction," international journal of advanced research in computer and communication engineering vol. 3, issue 9, september 2014

[5]. Heinrich W. L¨ollmann, Emre Yilmaz, Marco Jeub, and Peter Vary, "An Improved Algorithm for Blind Reverberation Time Estimation,"Institute of Communication Systems and Data Processing, RWTH Aachen University, 52056 Aachen, Germany

[6]. P. Bestagini, S. Milani, M. Tagliasacchi, and S. Tubaro, "Local tampering detection in video sequences, "in Proceedings of 15th IEEE International Workshop on Multimedia Signal Processing(MMSP 2013), Sept 2013, pp. 488493.

[7]. J. Haitsma and T. Kalker, "A Highly Robust Audio Fingerprinting System with an Efficient Search Strategy," Journal of New Music Research, vol. 32, no. 2, pp. 211–221, 2003.

[8]. S. Roy and Q. Sun, "Robust Hash for Detecting and Localizing audio video Tampering," in IEEE International Conference on Image Processing, S.Antonio, TX, 2007, vol. 6.