# Automatic detection of TV commercials

Brandon Satterwhite
and Oge Marques



TV advertising commercials are a critical marketing tool for many companies. Their interspersion within regular broadcast television programming can be entertaining, informing, annoying or a sales goldmine depending on one's viewpoint.

As a result, there are two major reasons for being able to detect commercial segments within television broadcasts. Interestingly, these two applications' goals—at least indirectly—are at odds with each other. One application seeks to identify and track when specific commercials are broadcast. Advertisers, in particular, like to verify that their contracts with broadcasters are fulfilled as promised. (The price of a commercial depends primarily on the popularity of the show it interrupts. The more people who are the product's demographics watching that program the better and usually the higher the cost, e.g. the Super Bowl. Thus, the advertiser wants to ensure that the commercial ran during the Super Bowl and not on a "That Girl" rerun at 4 am.)

The other group wants to detect commercials for the purpose of eliminating them from their recordings. This group is viewers who want to watch their recorded television shows without the annoyance of commercials. Video database maintainers would also appreciate the ability to automatically edit out commercials in stored shows and thereby decrease storage requirements. Of course, advertisers are strongly opposed to such devices because that defeats the purpose of the commercials. This article will discuss several algorithms that have been experimentally used to detect commercials, as well as devices that are currently available for this purpose.

## Characteristics of commercials

The problem of detecting commercials within television broadcasts is related to several—more general—problems in video processing. These issues include scene break detection, video segmentation, and video indexing and retrieval. However, commercial segments have certain characteristics that make them easier to identify than general video segments. These characteristics make it possible to use detection algorithms unsuitable for feature extraction from a general video database.

First, commercials are almost always grouped into blocks, typically consisting of four to 10 commercials each. As shown in Fig. 1, at the beginning and end of each commercial block and between each commercial in the block, several frames of monochrome black are displayed. On many stations, the observation has been that the last two to three commercials of a block are commercials promoting upcoming shows. Also, some countries (e.g. Germany) have laws requiring that every commercial block begin with a standard "commercial block introduction" sequence according to Lienhart et al.

Many television stations also have a practice of displaying a network logo in the corner of the screen during regular programming and then removing this logo during commercial breaks. Within a given television series, all episodes generally have commercial breaks scheduled at approximately the same time in the episode. Also, many commercials are repeated on a frequent basis, particularly for a given station.

Several other characteristics relate to the individual commercials. The duration of individual commercials is short, almost always less than ninety seconds, and typically it is an integer multiple of fifteen seconds. To capture viewers' attention in the small amount of time available to convey a message, commercials tend to be high in "action," typified by a high number of cuts between frames among other things. (Lienhart et al noted that the average "hard" cut rate in a sample of 200 commercials from German television was 20.9 cuts per minute, while the rate in the accompanying movie clips was only 3.7 cuts per minute.) There are usually a large number of frames with text containing the product or company's name. Also, to leave the product in the viewer's mind, the last few seconds of many ads consist of "still" shots of the product or slogan.

Other characteristics beyond the visual information are often present as well. The most noticeable characteristic, and the one most irritating to viewers, is the tendency of broadcasters to increase the volume level of the audio track during commercials. Another audio clue to the presence of commercials is that the delimiting black frames at the beginning and end of commercials are accompanied by silence in the audio track. Also, the dialogue on the audio track generally contains the product or company's name. Finally, when closed captioning is available for a television show, it is generally discontinued during commercial breaks.

No currently proposed detection algorithm utilizes all of these clues. Most algorithms, however, detect various combinations of them in order to improve detection rates.

## Detection schemes

There are two main categories of

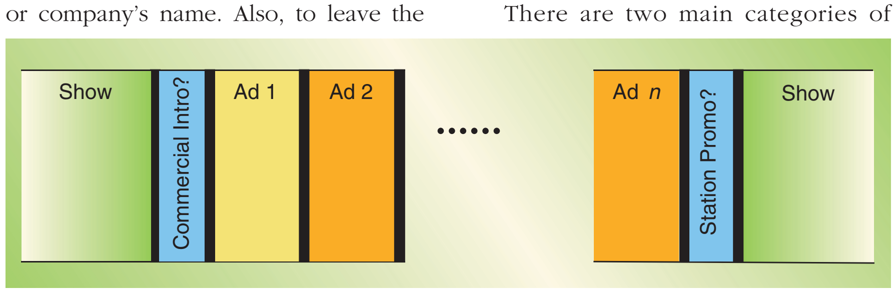


**Fig. 1 Structure of typical commercial block**

methods used to detect commercials. *Feature-based detection* relies on general characteristics of commercials to detect their presence. Any of the commercial characteristics mentioned earlier could be used to indicate the (possible) presence of a commercial. *Recognition-based detection* attempts to identify individual commercials in the broadcast as matching commercials it has already learned.

## Black frames and silences

The most common characteristics used in commercial detection are the delimiting black frames and silences. In locating black frames, the simplest method is to look at the average intensity value of the pixels in the image. The average intensity is determined easily in the analog domain [Hurst] and is the basis for most current commercial applications. The determination that a given frame is "black" is based on the average being below a pre-determined threshold value. Improved black frame detection can be accomplished by requiring that the standard deviation of the intensity values also be below a threshold according to Lienhart et al. Some work has also been done by Sadlier et al regarding a method to detect black frames in an MPEG-encoded bit stream, without the computational cost of decoding.

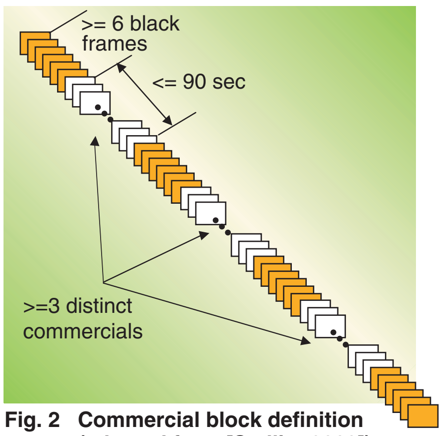


**Fig. 2 Commercial block definition (adapted from [Sadlier 2002])**

Silent audio frames may similarly be detected by examining the average volume level on the audio track. Most applications couple these two functions to decrease the likelihood of a false detection of a black frame or the detection of an irrelevant black frame within a program. By this rule, a black frame is only detected if it is accompanied by a silence. To further reduce the chance of a random black frame detection, most algorithms require that a certain number of consecutive black frames be detected together, usually five or six.

Once black frames can be reliably detected, the timing aspects of the commercial breaks can be exploited. Two black frame series detections may indicate a commercial segment is between them. Most algorithms establish a maximum time between black frame sequences for a segment to be considered a possible commercial. If the time between black frame sequences is greater than this, the segment is considered to be part of the program. The algorithm proposed by Sadlier et al sets this maximum commercial length at ninety seconds (2250 frames at 25 frames per second). This algorithm also looks at how many consecutive commercial segments occur to determine if the candidate commercial is in fact part of a commercial block. It requires that if a potential block does not contain at least three individual commercials, then it must be part of the program. (That is, if at least four black frame sequences occur with a maximum separation between each of ninety seconds, it is classified as a commercial break.)

Using 10 broadcast clips, Sadlier et al evaluated the algorithm (Fig. 2). The total amount of time was 315 minutes and included various genres such as sports, news and talk shows. The 10 broadcast clips contained a total of 11 commercial breaks as determined by human inspection. The algorithm detected all 11 commercial breaks, and none of the programming content was missed.

However, the algorithm did fail to detect parts of the last commercial in some of the blocks, incorrectly including them with the programming instead. Still, the algorithm performed reasonably well. Calculation of a performance measure called "recall," which is the percentage of commercial time correctly identified as such, showed that only one of the 11 clips had a recall rate below 85%. Eight of the clips had a recall rate greater than 98%.

## High cut rate and action

Another characteristic used in feature-based detection is the high cut rate typically observed in commercials. The problem of determining the cut rate of a video segment is basically the same as the problem of determining shot changes (where the video switches from one shot to another). Once shot changes have been located, determining the cut rate is merely a matter of counting.

A number of methods have been proposed to locate shot changes, most use statistics on differences in the color histogram from one frame to the next. Another method proposed by Wen et al uses a wavelet-based distance metric to quantify the difference between two frames and identify cuts.

One algorithm for using the cut rate in commercial detection, proposed by Lienhart et al, had two basic rules: 1) a candidate sequence must have a cut rate above five cuts per minute for its entirety, and 2) the cut rate must go above 30 cuts per minute at some point. This algorithm had a recall rate of 93.43% and a false detection rate of 0.09%, confirming the suitability of using strong hard cuts as a pre-filter for commercial blocks.

Some algorithms incorporate other editing techniques used frequently in commercials, such as fades and dissolves, to indicate the possible presence of commercials. Lienhart's group uses two additional metrics related to the high level of action in commercials. First, the "edge change ratio" describes the number of edge pixels (as found by an edge detection algorithm) entering and leaving a frame. The second metric, called motion vector length, describes the motion of objects in the image. It is similar to the motion vectors calculated in MPEG encoding. Detection methods based on these two metrics both had recall rates around 96% when used on their test database.

Naturally, feature-based detection is most effective when multiple characteristics are considered together. Lienhart et al created a combined system that has two steps. First, the black frame sequence detector and the cut-rate detector are used to find candidate commercial segments. Then those candidate segments are passed to the action detectors (edge change ratio and motion vector length) to find the exact commercial block limits. The advantage of this two-step system is that the more computationally expensive operations can be reserved for the second step.

## Recognition-based methods

Recognition-based detection methods are specialized video database systems that maintain a database of known commercials. To determine if the current segment of a television broadcast is a commercial, the segment is compared to known commercials using a query-by-example type operation. If a match is found, then the segment is almost certainly a commercial (depending on the precision of the matching algorithm).

Because of the computational expense involved in searching through a video database, most recognition-based algorithms use at least a simple feature-based detector—as a pre-selector—to determine candidate video segments, i.e. a shot segmentation algorithm or a black frame sequence detector. Their purpose is to determine the start point for the video segment to be sent to the database. Since the black frames or cuts are already being located for that purpose, it is convenient to look at their timing to perform a feature-based pre-selection.

Recognition-based systems are susceptible to problems in matching a segment from a broadcast to the same one in the database because of the variations caused by irregularities in the broadcast. Color levels of the same commercial can vary from station to station. Also, commercials are sometimes edited to shorten their length, which make them somewhat more difficult to match. Thus, any recognition-based system must be flexible enough in its search algorithm to allow for such variations. There is some evidence that, because of broadcasting variations, the color histogram techniques that are prevalent in video database indexing may not be ideally suited for recognizing commercials. The wavelet-based approach of Wen et al and the gradient method of Hampapur and Bolle are examples of algorithms that use non-color based indices to overcome this problem.

The recognition-based algorithm proposed by Lienhart et al uses a database-matching scheme that can match subsequences within video segments. This ability makes it possible to recognize edited commercials. This algorithm searches the database in two steps. The algorithm uses an index of color coherence vectors (CCV). These vectors are like color histograms but give some spatial information by indicating how many pixels are contained in "monochromatic" regions in the image.

As shown in Fig. 3, Lienhart et al used a sliding window to indicate the segment of the current broadcast to send to the database for a possible match. In the first step, a window of L seconds is compared to the first L+S seconds of the commercials in the database. If a potential match is found in the database, the comparison window is expanded to the full length of the candidate commercial. Because of the lower number of frames being compared, the first step in this algorithm is markedly shorter than a search using the entire commercial. This first step weeds out enough non-matches to provide a net decrease in computation time even though two searches are required to detect a single commercial. Experimentally, this algorithm correctly identified all 125 commercials in three hours of video when given a 200 commercial database in which to search. On average, the beginning and end frames of the commercials were detected to within 5 frames of the actual.

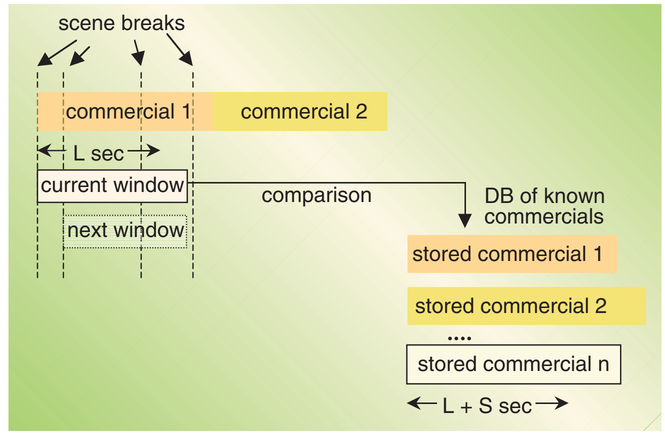


**Fig. 3 First step of recognition-based algorithm proposed by Lienhart et al. (adapted from [Lienhart 1997])**

Recognition-based systems face the drawback that commercials must be known (and indexed in the database) before they can be recognized in a broadcast. There are three possible modes of operation that have been proposed to accomplish this necessary commercial "learning." First, the user could indicate to the algorithm when a new commercial is encountered. The system would then store that new commercial in the database. Second, companies could compile databases of the most frequently aired commercials and sell such databases to users. The third, most useful, option is for the system to automatically learn new commercials as it encounters them.

Lienhart et al proposes a system for such automated commercial learning. It assumes that most commercials are already known. New commercials are entered into the database if they are surrounded by two previously known commercials (and are less than ninety seconds). Of course, this method will tend to miss ads such as station promos that generally appear at the end of commercial blocks.

One interesting area of possible future research—suggested by Del Bimbo and Colombo—is trying to use a recognition-based approach that recognizes not only pre-indexed commercials, but also commercials that have some semantic connection to previously learned commercials. Obviously, this area will benefit greatly from the ongoing research in the area of video databases.

## Applications

As noted, there are two major areas of application for commercial detection algorithms: "commercial trackers" and "commercial killers." Commercial trackers are designed to automatically audit the broadcast of commercials so advertisers can verify fulfillment of their "air play" contracts. Clearly, this application must use recognition-based methods because specific commercials are being sought out. If feature-based indicators are used within such recognition-based devices, it is desirable to adjust any threshold values to minimize false negatives. This way the chance of missing commercials will be minimal. (The corresponding increase in false positives—i.e. missed program time—will be counteracted with the recognition-based portion of the algorithm.)

Commercial killers try to remove commercials from the recordings so that viewers do not have to watch them on playback. Devices for this purpose started showing up in the mid-90s. Today, most major VCR brands offer an option, generally called "Commercial Advance," to do this. All the major brands rely on the same technology, which was developed by Iggulden in 1994. The algorithm is a simple one based on detecting black frame sequences and analyzing the timing between them. As a broadcast is recorded on the VCR, it keeps track of when the black frames occur. When the recording stops, it performs the necessary computations to determine where the commercial blocks are. This information is then encoded on the videotape. When the tape is subsequently viewed, the VCR automatically fast-forwards past commercial blocks.

The digital video recorder (DVR) made by ReplayTV uses a similar algorithm. According to their statistics, the algorithm has been observed to eliminate 96% of commercials under controlled test. They acknowledge that in real world use the success rate can be significantly lower, between 70% and 90%. Most DVR systems produced by other manufacturers offer a simple "Skip Forward" option that fast forwards by a fixed amount of time—generally set at 30 seconds—to allow the user to skip past commercials without actually having an automated system to detect them. This controversial feature has recently caused friction between DVR users and media companies' executives according to *USA Today*.

The use of a Commercial Advance option on DVRs is much more natural than on VCRs. First, DVRs have much more functionality than VCRs. Also, the digital nature of DVRs means the fast forwarding is instantaneous.

## Limitations

We see the research efforts reported in this article as early experimental works. A significant amount of additional effort will be required before they turn into robust commercial solutions. There are two main types of limitations:

- *Legal/commercial,* since developments in this field—particularly in the "commercial killers" category—will potentially aggravate the ongoing battles between content providers and consumers/viewers.
- *Technical,* since many of the proposed algorithms need to be enhanced and tested against much bigger and more diverse databases before being fully deployed and used on a large scale.

Advertisers and broadcasters have been pushed to change their methods—especially after the popularization of DVRs such as TiVo and their "skip commercials" capabilities—and have reacted in many different ways. Some TV executives threaten to raise cable and satellite costs as much as an additional $250 (USD) per year (see http://www.techtv.com/news/computing/story/0,24195,3391766,00.html) to make up for revenues lost to skipped ads. Others may be looking at how they can learn from studies on viewers' habits in terms of which ads they skip and which they don't. For example, the recent TiVo Commercial Viewing Report shows that users tend to skip commercials in comedies and general drama programs. But they will watch ads in reality TV, news and event programs. (see http://www.pcworld.com/news/article/0,aid,111015,00.asp). Yet, others claim that ads should be embedded in the main program itself (in a banner-like format).

Developments in this arena may lead to interesting, additional work in video analysis (e.g., automatic detection of program genre to activate/deactivate commercial skipping features, or automatic techniques for detection and removal of banner ads). With the switch to digital television devices, it should be much easier to implement the commercial detection algorithms discussed here in cost-efficient devices for home use. These devices also should perform much better.

Of course, advertisers need their commercials to be seen. Thus, broadcasters will want to make any possible programming changes to defeat such devices. So the end line on detection algorithms research keeps getting pushed forward.

## Read more about it


- C. Colombo, A. Del Bimbo, P. Pala. "Retrieval of Commercials by Semantic Content," *Multimedia Tools and Applications*, 13(1): 93-118, January 2001.
- A. Del Bimbo, P. Pala, L. Tanganelli. "Retrieval of Commercials based on Dynamics of Color Flows," *Journal of Visual Languages and Computing*, vol. 11, no. 3, June 2000.
- N. Dimitrova. "Digital Video for the Time Impaired," *IEEE Multimedia*, vol. 6, no. 2, pp. 14-17, 1999.
- A. Hampapur, R. Bolle. "Feature Based Indexing for Media Tracking, IEEE Int'l Conference on Multimedia and Expo, vol. 3, pp. 1709-1712, 2000.
- A.G. Hauptmann and M.J. Witbrock. Story segmentation and detection of commercials in broadcast news video," *ADL-98 Advances in Digital Libraries*, Santa Barbara, USA, 1998.
- R. Hurst. TV Commercial Editor, *Electronics Now*, vol. 65, no. 8, p.31, Aug. '94.
- J. Iggulden. Method and Apparatus for Controlling a Videotape Player to Automatically Scan Past Recorded Commercial Messages. US Patent Number 5333091. July 26, 1994.
- Y. Li, C. Kuo. "Detecting Commercial Breaks in Real TV Programs Based on Audiovisual Information," *Proceedings of the SPIE*, vol. 4210, pp. 225-236, 2000.
- R. Lienhart, C. Kuhmunch, W. Effelsberg. "On the Detection and Recognition of Television Commercials," *IEEE Conf. on Multimedia Computing and Systems*, pp. 509-516, 1997.
- ReplayTV, website, March 29, 2002.
- D. A. Sadlier, S. Marlow, N. O'Connor, N. Murphy. "Automatic TV Advertisement Detection from MPEG Bitstream," *WPRIS 2001 - International Conference on Enterprise Information Systems, Workshop on Pattern Recognition in Information Systems*, Setubal, Portugal , 7-10 July 2001.
- J.M. Sánchez, X. Binefa and J. Vitrià. "Shot Partitioning Based Recognition of TV Commercials," *Multimedia Tools and Applications*, 18(3), pp. 233-247, December 2002.
- J. M. Sanchez, X. Binefa. "AudiCom: A Video Analysis System for Auditing Commercial Broadcasts," *IEEE Int'l Conf. on Multimedia Computing Systems*, vol. 2, pp. 272-276, 1999.
- *USA Today*, 06 June 2002. "ReplayTV users sue for right to skip commercials." http://www.usatoday.com/tech/news/2002/06/06/replaytv-sue.htm
- X. Wen, T. D. Huffmire, H. H. Hu, A. Finkelstein. "Wavelet-based Video Indexing and Querying," *Multimedia Systems*, vol. 7, issue 5, pp. 350-358, 1999.


## About the authors


Brandon Satterwhite recently completed a Master's degree in Computer Science at Florida Atlantic University. He graduated summa cum laude with a B.S. in Mechanical Engineering at the University of Alabama. Currently, he is employed by Pratt & Whitney, a division of United Technologies, on the development of military jet engines.

Oge Marques is an Assistant Professor in the Department of Computer Science and Engineering at Florida Atlantic University in Boca Raton, Florida. He received his B.S. degree in Electrical Engineering from *Centro Federal de Educação Tecnológica do Paraná* (CEFET-PR) in Curitiba, Brazil, and a Master's degree in Electronic Engineering from *Philips International Institute of Technological Studies* in Eindhoven, The Netherlands, and a Ph.D. degree in Computer Engineering from Florida Atlantic University. Dr. Marques wrote (with Dr. Borko Furht) the book, *Content-Based Image and Video Retrieval* (Kluwer Academic Publishers, Boston, MA, 2002) and is also one of the editors-in-chief of *The Handbook of Video Databases* (CRC Press, Boca Raton, FL, 2004).