



HAL
open science

Unsupervised Credit Detection in TV Broadcast Streams

Yannick Benezeth, Sid-Ahmed Berrani

► **To cite this version:**

Yannick Benezeth, Sid-Ahmed Berrani. Unsupervised Credit Detection in TV Broadcast Streams. International Symposium on Multimedia, Dec 2010, Taichung, Taiwan. pp.175 - 182, 10.1109/ISM.2010.33 . inria-00545500

HAL Id: inria-00545500

<https://inria.hal.science/inria-00545500>

Submitted on 16 Oct 2012

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Unsupervised Credit Detection in TV Broadcast Streams

Yannick Benezeth and Sid-Ahmed Berrani

Orange Labs - France Telecom R&D

4, rue du Clos Courtel - BP 91226

35512 Cesson-Sévigné Cedex - France

Email: {yannick.benezeth,sidahmed.berrani}@orange-ftgroup.com

Abstract—This paper proposes an unsupervised method for detecting credits in TV streams. Identifying credits in TV streams allows to precisely determining boundaries of TV programs and hence, to extract specific and high valuable TV programs. The proposed detection solution is based on the temporal stability of opening and closing credits. Consequently, from a linear TV stream, we detect sequences that are broadcasted several times on a stable schedule with a clustering-based approach. These repeated sequences include opening and closing credits but also commercials, trailers etc. In order to select, among repeated sequences, those which are effectively credits, their temporal stability and their metadata consistency are analyzed. Since recurring programs are usually broadcasted at the same time(s) of the day, the temporal distribution of occurrences of each repeated sequence is studied. The Electronic Program Guide (EPG) is then used to validate the selected sequences and to distinguish between opening and closing credits. This method is entirely unsupervised and no prior information is required. Extensive experimental results validating our approach are presented.

Keywords-Credit detection, Repeated sequence detection, Hierarchical clustering, TV stream structuring.

I. INTRODUCTION

We have witnessed, in recent years, a rapid increase of the broadcasted TV streams number. Consequently, the amount of digital video available to viewers is becoming extremely large. Paradoxically, available tools for interacting with this multimedia content are still limited. With the purpose of building novel TV services such as TV-on-demand or catch-up TV, the programs need to be extracted from the TV stream, indexed and stored. Unfortunately, the metadata broadcasted with the TV stream, namely the EPG (Electronic Program Guide) or the EIT (Event Information Table), are approximate and incomplete [1], [2]. They do not enable an automatic and reliable extraction of programs from the stream. On the other hand, manual processing is clearly unfeasible regarding the huge amount of streams to process. Therefore, automatic tools for precisely extracting programs are highly relevant and required for novel TV services.

Few techniques have been proposed to automatically segment and index TV streams. Many previous works focus on shots partitioning (e.g. [3], [4]), scenes organization [5] and few ones attempt to segment the TV stream in programs. These last methods attempt to directly detect programs (e.g. [6]) or can detect inter-programs (e.g. [7]).

Although TV programs are various and heterogeneous, they usually share a common structure: an opening credit, program content followed by a closing credit. Moreover, we empirically observed that a generalist TV channel usually broadcasts more than 80% of recurring programs (daily or weekly ones). Furthermore, recurring programs are transmitted on a stable schedule in order to secure the loyalty of viewers. In other words, a recurring program is usually broadcasted periodically (daily or weekly) at the same time.

The work presented in this paper is based on the above observations. We claim that the schedule stability of recurring programs can be used to automatically and reliably detect opening and closing credits. In the following, we will consider that opening credits are the sequences that appear at the beginning of a TV program. Opening credits often list the most important members of the production and may be composed of text, music and a real or synthetic video clip. Closing credits appear at the end of the TV program and list the production crew and actors. Closing credits may be composed of text and music. Examples of opening and closing credits are presented in Figure 1.

Identifying opening and closing credits significantly contribute to precisely determine the boundaries of TV programs. In the following, we call TV stream macro-segmentation the process which consists in determining the start and end of each TV program in the stream. Since recurring programs represent 80% of broadcasted programs, the whole macro-segmentation process should not be solely based on credit detection. Nevertheless, the macro-segmentation quality can be greatly improved when the opening and closing credits are reliably and precisely detected. Indeed, the macro-segmentation quality depends on the difference between the detected start (resp. end) time and the actual start (resp. end) time. In order to measure how much the performance of TV stream macro-segmentation is improved with the credit detection, we present in section VI experimental results of TV streams macro-segmentation with and without considering the proposed credit detection solution.

Besides improving the macro-segmentation quality, identifying online opening credits sequences can be used to detect specific programs as soon as they are broadcasted in the live stream. Recommendation systems, informing users that

a program of interest is starting on another channel, can be designed relying on this technology.

The main advantages of our method are listed below:

- 1) In order to reduce manual processing, we designed a fully automatic and unsupervised approach which extracts credits of recurring programs from the raw TV stream.
- 2) No assumption is formulated on the audio-visual content of credits (e.g. the presence of text, music etc.).
- 3) Using the EPG, we can distinguish between opening and closing credits.
- 4) The proposed method is able to deal with different recurring frequencies (daily, weekly...), and efficiently handles late modifications of the schedule with respect to the expected one. Indeed, once a video sequence is associated with a credit, it can be detected whenever it appears.

The rest of the paper is organized as follows. Related works are first presented in Section II. Then, the method for repeated sequence detection is presented in Section IV, followed by the temporal stability and the EPG consistency analysis in Section V. Finally, Section VI presents experimental results validating our approach. Section VII concludes the paper and outlines future works.

II. RELATED WORKS

To the best of our knowledge, only few papers have directly addressed the problem of credit detection. Conversely, the literature is abundant concerning related topics such as commercial detection (e.g. [8]), the identification of events [9] or the structuring of specific TV programs such as news [10] or sport games [11]. Most of these methods are based on the visual signal analysis, some focus on the audio signal analysis [12] or on the combination of both [13].

Methods related to credit detection can be divided into three categories, namely (1) the supervised approaches, (2) the feature-based approaches and (3) the TV stream structure based approaches. Methods in (1) look forward at recovering sequences manually introduced in a reference database. These techniques stand on a content-based matching method such as fingerprinting [14] or perceptual hashing [15]. The main drawback of this kind of methods is that the database has to be manually created for each TV channel. And thus, the flexibility is degraded while the maintenance is laborious.

Then, the feature-based approaches, methods in (2), attempt to detect credits based on their intrinsic features. For example, Li et al. [16] present a method for detecting and tracking text in video. As text is an important part of closing credits (and sometimes of opening ones), this feature is interesting to study. The main drawback of this approach is the costly processing time required to analyze each frame. For example, authors explain that it takes about 1 second to track text blocks on each frame for the

closing credits of the movie Star Wars. Then, Naphade and Huang [12] propose to use Hidden Markov Models to model the dynamics of the soundtrack and detect audio events. Even if authors do not explicitly mention it, their method can be applied to the credit detection task. Indeed, opening credits of entertainment TV shows are often composed of music and applause followed by the announcer speech. This sequence of soundtrack states (music, applause, speech) can help in detecting opening credits. Then, Wang et al. [13] propose a multimodal scheme for program segmentation. Their method is based on visual, auditory and textual information. Their visual features, called *POIM* for program-oriented informative images, attempt to characterize opening and closing credits. The hypothesis is that credits share a set of common features (mainly text with synthetic images) to be detectable with low-level features. Nevertheless, the intra-class variability of opening and closing credits is very high on a generalist TV channel because broadcasted programs are very heterogeneous. Figure 1 illustrates the high intra-class variability of opening and closing credits.

Finally, methods in (3) use the structure of TV streams to detect credits. The work presented by Liang et al. [17] is based on the schedule stability of credits to detect programs. During an offline training stage, the keyframes of each shot is compared with the keyframes of the shots at the same time on previous days. If shots are identical, this shot is saved in a program boundaries database. Programs are then detected using these learned programs boundaries (under the constraint of time of appearance). This approach suffers from some weaknesses. First, as metadata are not used, the difference between credits and a repeated commercial and the difference between opening and closing credits are unmanageable. Moreover, Liang et al. [17] define a program as a temporal segment between two detected sequences. Consequently, when a credit is not detected, a shift will appear in the results (between programs and inter-programs sequences). This approach cannot deal with complex TV streams and cannot take into account changes of TV schedule w.r.t. the expected one.

In this paper, we propose to overcome the limitations of the previous approaches. We present a fully automatic and unsupervised method that handles different TV program recurring frequencies and late changes of the schedule. Furthermore, using metadata, our method is able to distinguish between opening and closing credits.

III. METHOD OVERVIEW

The proposed credit detection method is based on repeated sequence detection. Consequently, from a linear TV stream, we aim to detect sequences that are broadcasted several times. This repeated sequence detection is performed on an accumulated amount of TV stream. We present in Figure 2 two possible processing schemes: periodical or on demand



Figure 1. Examples of opening and closing credits illustrating their high intra-class variability.

credit detection. In Figure 2.(a), the credit detection process is periodical (period T in the figure). First, the system needs to accumulate a sufficient amount of stream to launch the first periodical credit detection (at $t = 3T$ in the figure). The analysis is performed on the previously accumulated stream. Then, the system starts accumulating the stream again and a new credit detection can be performed at the next period ($t = 4T$). In Figure 2.(b), the credit detection process is on demand. One can be interested in the detection of credits for some specific interval (e.g. one day) and the credit detection is performed on the specified interval plus the required accumulated amount of video stream.

Intuitively, one can note that the amount of accumulated TV stream depend on the recurring frequency of TV programs. For example, if one is interested in detecting credits of daily recurring programs, an accumulated TV stream of 3 or 4 days is sufficient. At the opposite, if one is interested in detecting credits of weekly recurring programs, obviously, the accumulated amount of TV stream should be at least of eight days. An analysis on this accumulated amount is presented in Section VI-C.

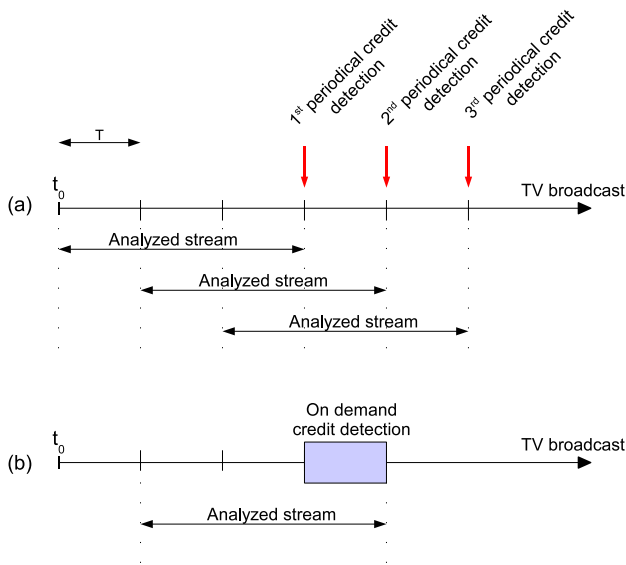


Figure 2. Processing scheme of the credit detection method. (a) Periodical credit detection. (b) On demand credit detection.

The detected repeated sequences include credits but also others video sequences such as trailers, commercials, etc. Recurring programs are usually broadcasted roughly at the same time(s) of the day. Consequently, in order to select among repeated sequences those which are effectively credits, the temporal distribution of occurrences of each repeated sequence is studied. The metadata is then used to validate the selected sequences and to distinguish between opening and closing credits. Details are given in the following parts.

IV. REPEATED SEQUENCE DETECTION

The first step of our credit detection method is the detection of repeated sequences. These repeated sequences are near identical sequences which are broadcasted at least twice. The detection is performed on an accumulated amount of TV stream (from 3 days to one week, cf. Figure 2). A repeated sequence is composed of a set of occurrences which correspond to each broadcast of the repeated sequence.

We use the technique described in [18] to detect repeated sequences. First, low-level visual descriptors of the video are extracted. For each frame, a 64-bits DCT-based descriptor, called *BVD*, is computed. Then, a DCT-based 30-dimensional descriptor, called *KVD*, is processed to carefully describe chosen keyframes. *KVDs* are more robust than *BVDs*. They are used to group similar keyframes. The number of *KVDs* per cluster ranges from two to several dozens. It corresponds to the number of times keyframes are repeated.

Once the clusters are detected, they are analyzed in order to create the repeated sequences. First, clusters that contain *KVDs* from the same shot or from neighboring shots are removed. Then, a set of measures are used. These measures are based on the temporal diversity of *KVDs* within a cluster and on the inter-cluster relationship (if *KVDs* of two clusters are interlaced). The repeated sequences boundaries are determined by extending created occurrences with a frame-based matching procedure using the *BVDs*. For more details, the reader can refer to [18].

V. CREDIT DETECTION

The detected repeated sequences include opening and closing credits but also commercials, trailers etc. In order to select, among repeated sequences, those which are effectively credits, the following two criteria are used:

- 1) the temporal stability,
- 2) programs consistency (using the EPG).

Since recurring programs are usually broadcasted at the same time(s) of the day, the temporal distribution of the occurrences of each repeated sequence is first studied. Then, the EPG is used to validate the selected sequences and to distinguish between opening and closing credits. These steps are described in this section.

A. Temporal stability analysis

The first selection step is performed using the temporal stability of occurrences of each repeated sequence. In the following, the reasoning is based on the hypothesis that recurring programs are usually broadcasted at the same time(s) of the day. Consequently, occurrences of each repeated sequence are projected onto their time of appearance and their temporal distribution is then studied. In Figure 3, we present the temporal distribution (along 24 hours) of occurrences of three repeated sequences. The occurrences of the repeated sequence presented in the timeline (a) are randomly scattered along the day while the occurrences of the repeated sequences shown in (b) and (c) are concentrated around one or two specific schedules. The occurrences of a repeated sequence that is a credit are likely to be densely distributed around some specific times. In the example of Figure 3, the repeated sequence (b) is a credit of a recurring program that repeats once a day and the repeated sequence (c) is a credit of a program that repeats twice a day.

Then, with the intention to analyze the temporal distribution of the occurrences of each repeated sequence, a clustering is performed to aggregate contiguous occurrences. The correspondence between the repeated sequence and one or more TV program in the EPG have not been established yet. Consequently, the recurring frequency of the concerning program is a priori not known. So, the number of expected clusters is also not known. Therefore, an unsupervised clustering method is performed. We use a bottom-up agglomerative hierarchical clustering [19] with the average linkage strategy. The distance between occurrences is the difference in time between their times of appearance. Each repeated sequence is analyzed alone. Its occurrences are clustered and a hierarchical structure convenient for further analysis is built. We present in Figure 4 an example of a hierarchical structure, called dendrogram, of the occurrences of a repeated sequence.

In order to decide if the occurrences of a repeated sequence are credits or not, we analyze the density of each node of the associated hierarchical structure. The lower the temporal variance of the occurrences of a node is, the more densely its population (the occurrences) will be distributed around one specific time. To keep only densely populated clusters, if a node has a temporal variance value above an empirically determined threshold, we cut edges gathering

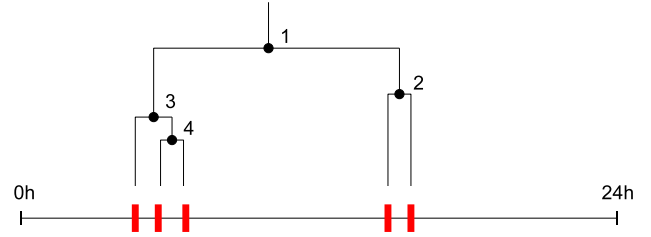


Figure 4. Dendrogram illustrating the occurrences organization of a repeated sequence resulting from the hierarchical clustering.

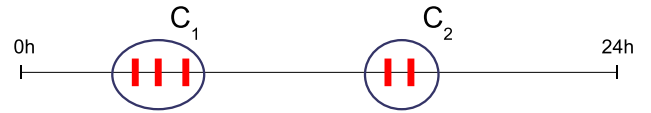


Figure 5. Example of clustering result illustrating two subsets composed of two and three occurrences.

its children. For example, in the dendrogram of Figure 4, the edge gathering the children of node 1 is cut. The nodes 2 and 3 are densely populated and are consequently maintained. We present in the Figure 5, the clustering result obtained from the dendrogram of Figure 4.

This allows us to deduce the clusters associated with a repeated sequence. They are analyzed in order to decide if the occurrences of the repeated sequence are credits. Few rules are applied.

First, if one of the clusters is a singleton, this repeated sequence is rejected. Its occurrences are not credits. This criterion can be moderated regarding the repeated sequences detection quality and the false detections tolerance. However, we have observed in our experiments that a severe decision rule leads to the best result.

Second, if the sequence is repeated more than n times a day, that is, if the number of generated clusters is greater than n , the sequence is rejected. The maximum number of times a sequence can be repeated per day has to be fixed w.r.t. the considered channels. Values of 2 or 3 are convenient for most of generalist TV channels. In our experiments, we have chosen $n = 2$.

B. Program consistency analysis

The temporal distribution analysis is not sufficient to reject *all* the repeated sequences that are not credits. Some commercials or trailers can be broadcasted on a relative stable schedule. In order to filter out these *false alarms*, we use the EPG and we study neighboring programs. The objective is to select repeated sequences that are always broadcasted before or after the same program. Doing that, we are also able to distinguish between opening and closing credits. The EPG provides basic information on the channel schedule, such as the approximate start and end times, the title and sometimes a synopsis of programs.

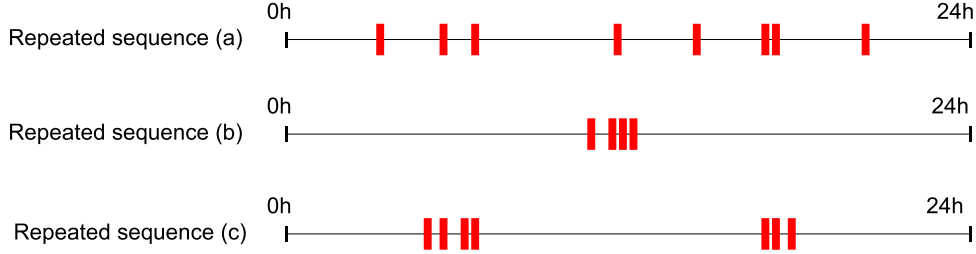


Figure 3. Occurrences distribution of repeated sequences onto their time of appearance. (a) A commercial with occurrences randomly scatter along the day. (b) A credit of a recurring program that repeats once a day. (c) A credit of a program that repeats twice a day.

For each repeated sequence that is selected after applying the temporal stability analysis step, the title of the preceding and the following program is determined for each occurrence. If both the preceding and the following programs are inconsistent for all of the occurrences of the repeated sequence, then the repeated sequence is rejected. It is worth pointing out here that the imprecision of the start and end times of the EPG is not a problem. What is important is only the order of programs.

Besides the metadata consistency analysis, we also measure the temporal distance between the occurrences of a repeated sequence and the boundary of the corresponding program in the metadata. If this average distance is beyond a threshold, the repeated sequence is rejected. In the example of Figure 6, because the program *News* is consistent with all occurrences of the repeated sequence, the average distance is computed between the occurrences and the start of the program *News*.

Obviously, knowing the position of the detected credits and the associated program, we are able to distinguish between the opening and closing credits. We present in Figure 6 an example of an opening credit in which the program following each occurrences of a repeated sequence is always *News*, whereas the preceding ones are different. One can note that, in the day 2, the occurrence is located in the *Inside Africa* program due to the metadata imprecision but with the consistency of the following program, we are able to label the repeated sequence as the opening credit of the recurring *News* program.

VI. EXPERIMENTAL RESULTS

We present in this section experimental results validating our approach. We first present the credit detection performance on two French generalist TV channels. Then, we present the benefits of the credit detection on the TV stream macro-segmentation performance and also the influence of the amount of accumulated TV stream.

In the following, we denote **DS1** the dataset collected from the channel TF1 and **DS2** the dataset collected from the channel France2. Both DS1 and DS2 last five days each. In order to automatically measure the performance of our

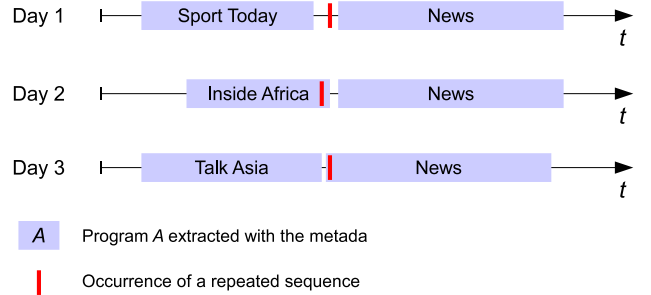


Figure 6. Opening credit example in which the following program is consistent for all the occurrences of the associated repeated sequence.

approach, credits of recurring programs in both datasets have been manually delimited.

A. Credit detection performance

The evaluation of the credit detection performance has been performed on DS1 and DS2. We have restricted the analysis to daytime programs (from 8:00 am to 23:59 pm) because nighttime ones are mainly rebroadcasts.

In the following, each broadcast of the same recurring program is denoted as "episode".

The Table I describes our datasets. We can observe that all opening credits and most of closing credits are stable, i.e. identical for all the episodes. Few recurring programs have a different closing credit for each episode. Consequently, these credits cannot be detected with a repeated sequences-based approach. They have been excluded from our evaluations.

Table I
STABILITY OF RECURRING PROGRAMS CREDITS.

	# of recurring programs	# of stable credits	
		Opening	Closing
DS1	10	10	6
DS2	14	14	7

We have applied our solution on DS1 and DS2. The obtained results have been evaluated using the precision and recall measures. They are summarized in Table II.

First, we can observe that most of the opening credits have been detected without any false detection. Actually, only three opening credits have not been detected. These misdetections are due to the repeated sequence detection method which has not detected all of the repeated sequences. Second, all closing credits have been detected with a reasonable number of false positives (i.e. a precision value of 0.87 for DS1 and 0.77 for DS2). These false detections correspond to the small video sequences announcing the next program and that share most of closing credits properties.

Table II
CREDIT DETECTION RESULTS.

		Precision	Recall
DS1	Opening Credits	1.0	0.80
	Closing Credits	0.86	1.0
DS2	Opening Credits	1.0	0.92
	Closing Credits	0.77	1.0

B. TV stream macro-segmentation improvement using credits detection

As explained in the section I, about 80% of broadcasted programs are recurring ones. We claim that the credit detection method is able to greatly improve the TV stream macro-segmentation process. We recall that the goal of the TV stream macro-segmentation is to determine precisely the start and the end times of each TV program in the stream. We use in this experiment the TV stream macro-segmentation method proposed in [7]. Although the goal is to detect program boundaries, this system focuses on detecting inter-programs (commercials, trailers, etc.). The main reason is that it is easier to detect inter-programs than programs. Programs do not share any common features. On the other hand, inter-programs exhibit detectable characteristics. Since the stream is only composed of programs and inter-programs, programs boundaries are given indirectly by inter-programs. The inter-program detection is based on the detection of separations between commercials, on the detection of repeated sequences and on the recognition of previously detected inter-programs.

In this experiment, program boundaries detected with the Naturel et al. [7] method are replaced with the boundaries of the detected credits. For example, if the boundaries of a program have been extracted by the method of Naturel et al., and one or both of the credits of the same program have been detected by our technique, then the boundaries of the

programs are modified to those derived from the detected credits.

We present in Tables III and IV quantitative results illustrating the benefits of using the credits detection to improve the accuracy of the TV stream macro-segmentation. This experiment has been performed on the dataset DS1. The results are presented distinguishing the TV series and other TV programs such as news, entertainments etc. We first present in Table III the proportions of programs for which the start (resp. end) time is less than 1 second away from the actual start (resp. end) time given by the ground truth. This measure indicates almost perfect segmentation. We can observe that the proportion of correctly detected programs has increased with the credit-based refinement step. The proportion of perfect detection of program starts has been improved to 100%. Nevertheless, we can observe that the percentage of perfectly segmented TV programs ends has slightly decreased. This is due to a trailer that has been detected as a credit. This trailer is systematically broadcasted after each episode of a TV program and has been added to the extracted episodes. This trailer lasts however only five seconds.

Table III
PROPORTION OF PERFECTLY SEGMENTED PROGRAMS WITH AND WITHOUT USING THE CREDIT-BASED REFINEMENT STEP.

	Naturel et al. [7]		Naturel et al. [7] + detected credits	
	Start	End	Start	End
TV Series	100%	87%	100%	94%
Others	91%	96%	100%	86%

Then, we present in Table IV the macro-segmentation accuracy. It shows the macro-segmentation imprecision of the detected programs start and end times. The imprecision is the absolute value of the difference between the detected start (resp. end) time w.r.t. the accurate start (resp. end) time. Table IV compares the accuracy obtained with the Naturel et al. [7] method and their method refined with the detected credits by our technique. The imprecision of the start time has been decreased to 0 second with the credits-based refinement while the end time imprecision concerning TV series has been decreased from 34 seconds to 3 seconds. We can also note that despite the trailer wrongly detected as a credit, the end time imprecision value has decreased from 34 seconds to 18 seconds.

C. Analysis on the amount of accumulated TV stream

The analysis of the accumulated stream amount influence on the credit detection performance has been performed on DS1. We present in Table V the credit detection performance

Table IV
IMPACT OF CREDIT DETECTION ON THE IMPRECISION OF
MACRO-SEGMENTATION.

	Naturel et al. [7]		Naturel et al. [7] + detected credits	
	Start imprecision	End imprecision	Start imprecision	End imprecision
TV Series	0s	34s	0s	3s
Others	12s	34s	0s	18s

varying the amount of accumulated stream on which the repeated sequence detection is performed.

Table V
ANALYSIS OF THE AMOUNT OF ACCUMULATED TV STREAM.

	Opening credits		Closing credits	
	Precision	Recall	Precision	recall
2 days	0.50	0.20	0.25	0.16
3 days	1.00	0.20	0.33	0.16
4 days	1.00	0.60	0.75	0.33
5 days	1.00	0.80	0.86	1.00
6 days	1.00	0.80	0.62	1.00
7 days	1.00	0.80	0.62	1.00
8 days	0.88	0.72	0.71	0.71
9 days	0.88	0.57	0.71	0.71
10 days	0.88	0.50	0.75	0.75

First, the Precision/Recall measures trend changes from the eighth day. From 2 days to 7 days, only credits of daily recurring programs are considered. Then, from the eighth day, credits of weekly recurring programs are also considered. The detection of weekly recurring program credits is challenging, a video sequence that lasts few seconds and that repeats only twice, must be detected in a 8 days or more video length. Consequently, we can observe that results deteriorate from the eighth day.

Second, because the number of missed credits decreases when the available amount of video increases, we can observe that the Recall measures increase from 2 days to 7 days. Nevertheless, the number of false detection starts increasing from the sixth day. Consequently, the optimal amount of accumulated stream to perform the credit detection is 5 days.

VII. CONCLUSIONS AND FUTURE WORKS

In this paper, we have proposed a method for detecting credits in TV streams. The presented method is entirely unsupervised and can extract credits of recurring programs from the raw TV stream. The proposed detection scheme

is based on the temporal stability of opening and closing credits. From a linear TV stream, we detect sequences that are broadcasted several times on a stable schedule. Then, their metadata consistency is analyzed in order to select, among the repeated sequences, those which are effectively credits.

We have shown in the experimental part that the proposed method is highly relevant for improving the TV stream macro-segmentation accuracy.

Future works will focus upon the integration of the credits detection within the TV stream macro-segmentation process. Particular attention will be given to the updating process of the credit dataset. Then, the possibility of using other information such as audio signal or subtitles will be studied. And finally, the metadata imprecision can lead to errors with our current method in the case of short programs (that last few minutes). Future works will focus on a better management of this imprecision in order to handle the detection of short programs credits.

REFERENCES

- [1] S.-A. Berrani, P. Lechat, and G. Manson, "TV broadcast macro-segmentation: Metadata-based vs. content-based approaches," *Proc. of the ACM International Conference on Image and Video Retrieval*, pp. 325–332, July 2007.
- [2] J.-P. Poli and J. Carrive, "Modeling television schedules for television stream structuring," in *Proc. of the International Multimedia Modeling Conference*, Singapore, January 2007, pp. 680–689.
- [3] N. Patel and I. Sethi, "Video shot detection and characterization for video databases," *Pattern Recognition*, vol. 30, no. 4, pp. 583–592, 1997.
- [4] J. Ren, J. Jiang, and J. Chen, "Shot boundary detection in mpeg videos using local and global indicators," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 8, pp. 1234–1238, 2009.
- [5] Z. Rasheed and M. Shah, "Scene detection in hollywood movies and TV shows," in *Proc. of the IEEE international Conference on Computer Vision and Pattern Recognition*, Madison, USA, June 2003, pp. 343–351.
- [6] E. El-Khoury, C. Senac, and P. Joly, "Unsupervised TV program boundaries detection based on audiovisual features," in *Proc. of the International Conference on Visual Information Engineering*, Xi'an, China, July 2008, pp. 498–503.
- [7] X. Naturel and S.-A. Berrani, "Content-based TV stream analysis techniques toward building a catch-up TV service," in *Proc. of the IEEE International Symposium on Multimedia*, San Diego, USA, December 2009, pp. 412–417.
- [8] L. Duan, J. Wang, Y. Zheng, J. Jin, H. Lu, and C. Xu, "Segmentation, categorization, and identification of commercial clips from TV streams using multimodal analysis," in *Proc. of the ACM international conference on Multimedia*, Santa Barbara, USA, October 2006, pp. 201 – 210.

- [9] D. Zhong and S. Chang, "Real-time view recognition and event detection for sports video," *Journal of Visual Communication and Image Representation*, vol. 15, no. 3, pp. 330–347, 2004.
- [10] X. Gao and X. Tang, "Unsupervised video-shot segmentation and model-free anchorperson detection for news video story parsing," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 12, no. 9, pp. 765–776, 2002.
- [11] L. Xie, P. Xu, S. Chang, A. Divakaran, and H. Sun, "Structure analysis of soccer video with domain knowledge and hidden markov models," *Pattern Recognition Letters*, vol. 25, no. 7, pp. 767–775, 2004.
- [12] M. Naphade and T. Huang, "Stochastic modeling of soundtrack for efficient segmentation and indexing of video," in *Proc. of the Conference on Storage and Retrieval for Multimedia Databases*, San Jose, USA, January 2000, pp. 168–176.
- [13] J. Wang, L. Duan, Q. Liu, H. Lu, and J. Jin, "A multi-modal scheme for program segmentation and representation in broadcast video streams," *IEEE Transactions on Multimedia*, vol. 10, no. 3, pp. 393–408, 2008.
- [14] A. Joly, C. Frelicot, and O. Buisson, "Robust content-based video copy identification in a large reference database," in *Proc. of the International conference on Image and Video Retrieval*, Urbana-Champaign, USA, July 2003, pp. 511–516.
- [15] C. De Roover, C. De Vleeschouwer, F. Lefebvre, and B. Macq, "Robust video hashing based on radial projections of key frames," *IEEE Transactions on Signal processing*, vol. 53, no. 10, pp. 4020–4037, 2005.
- [16] H. Li, D. Doermann, and O. Kia, "Automatic text detection and tracking in digital video," *IEEE Transactions on Image Processing*, vol. 9, no. 1, pp. 147–156, 2000.
- [17] L. Liang, H. Lu, X. Xue, and Y. Tan, "Program segmentation for TV videos," in *Proc. of the IEEE International Symposium on Circuits and Systems*, Kobe, Japan, may 2005, pp. 1549–1552.
- [18] S.-A. Berrani, G. Manson, and P. Lechat, "A non-supervised approach for repeated sequence detection in TV broadcast streams," *Signal Processing: Image Communication, special issue on "Semantic Analysis for Interactive Multimedia Services"*, vol. 23, no. 7, pp. 525–537, 2008.
- [19] A. Jain, M. Murty, and P. Flynn, "Data clustering: a review," *ACM computing surveys*, vol. 31, no. 3, pp. 264–323, 1999.