

## Review Article

# Systematic Review of Multiple Contents Synchronization in Interactive Television Scenario

**Ricardo Mendes Costa Segundo and Celso Alberto Saibel Santos**

*Federal University of Espírito Santo, Campus Universitário de Goiabeiras, Avenida Fernando Ferrari, S/N Goiabeiras, 29060-970 Vitória, ES, Brazil*

Correspondence should be addressed to Celso Alberto Saibel Santos; [saibel@inf.ufes.br](mailto:saibel@inf.ufes.br)

Received 20 November 2013; Accepted 3 February 2014; Published 19 March 2014

Academic Editors: G. Giambene and S. Rapuano

Copyright © 2014 R. M. Costa Segundo and C. A. Saibel Santos. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

*Context.* Interactive TV has not reached yet its full potential. How to make the use of interactivity in television content viable and attractive is something in evolution that can be seen with the popularization of new approaches as the use of second screen as interactive platform. *Objective.* This study aims at surveying existing research on Multiple Contents TV Synchronization in order to synthesize their results, classify works with common points, and identify needs for future research. *Method.* This paper reports the results of a systematic literature review and mapping study on TV Multiple Contents Synchronization published until middle 2013. As result, a set of 68 papers was generated and analyzed considering general information such as sources and time of publication; covered research topics; and synchronization aspects such as methods, channels, and precision. *Results.* Based on the obtained data, the paper provides a high level overview of the analyzed works; a detailed exploration of each used and proposed technique and its applications; and a discussion and proposal of a scenario overview and classification scheme based on the extracted data.

## 1. Introduction

Multimedia systems allow the data streams integration of different types, including continuous (audio and video) and discrete media (text, data, and images). Synchronization is essential for the integration of these media [1] and is focus of researches for a long time [2, 3]. Most of these works use a common taxonomy proposed by Cesar and Chorionopoulos [4] to classify multimedia synchronization.

This classification [5] is based on multimedia abstraction layers (Figure 1): in the media layer an application operates on a single continuous media stream, which is treated as a sequence of LDUs/MDUs (logical Data Units/Media Data Units); the stream layer allows the application to operate on continuous media streams as well as on groups of media streams; the object layer allows for simpler and exact specification of playout sequences, where each media object relates to a time axis and defines a sequence of events.

In the media layer, the intrastream synchronization deals with the maintenance, during the playout, of the temporal relationship within each time-dependent media stream, that

is, between the MDUs of the same stream. In the stream layer the interstream synchronization refers to the synchronization, during the playout, of the playout processes of different media streams involved in the application and live synchronization deals with the presentation of information in the same temporal relationship as it was originally collected. The object layer presents synthetic synchronization where various pieces of information (media objects), at presentation time, must be properly ordered and synchronized in space and time [5].

Previous classification, however, does not consider the problem of synchronizing media streams across multiple separated locations, which can be found in literature as multipoint [6], group [7], or Inter-Destination Multimedia Synchronization (IDMS) [8]. This synchronization level is on top of the object layer and should be presented in what is called semantic layer (Figure 2). The semantic layer allows communication, search, retrieval, and interpretation of playouts and its contents. Besides IDMS, the semantic layer also deals with context synchronization (cross media, mash-ups, etc.). It considers that some authors use synchronization

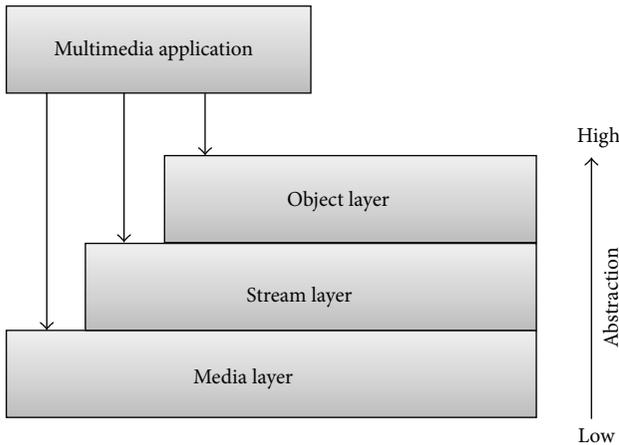


FIGURE 1: Synchronization abstraction layers—[5].

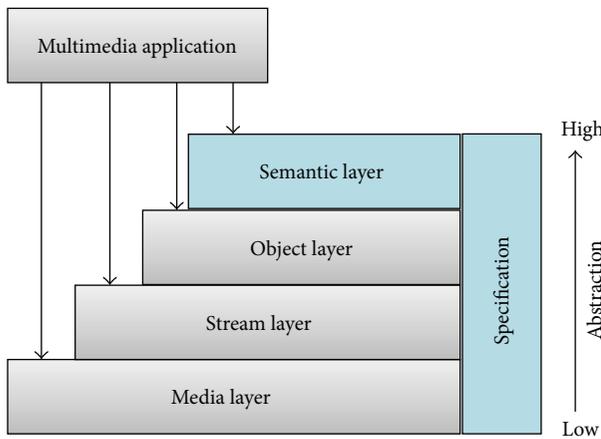


FIGURE 2: Abstraction layers with semantic layer.

in multimedia systems in a more general and widely used sense as comprising content, spatial, and temporal relations between media objects. In Figure 2, specification layer is also considered in the model proposed by Blakowski and Steinmetz [9]. The specification layer is an open layer that contains applications and tools that allow one to create synchronization specifications.

In the model derived from Meyer's one, the specification is not considered an isolated layer but one that is bound to all layers once every layer needs its own specification.

The following case shows that the use of only the three layers may not be sufficient to provide a satisfactory experience to TV viewers.

Figure 3 presents a Digital TV Application being executed on Brazilian Ginga middleware [10]. All media units are being correctly played (media layer); both video and audio are synchronized (stream layer) and the application media and video are correctly positioned and all relations defined by the NCL document are working just as defined (object layer). All three original layers specifications are being accomplished, but the user will notice something wrong at an upper level. While the video, audio, and EPG provides information about



FIGURE 3: Semantic layer specification problem.

a soccer match, the multimedia content of the DTV application transmitted “synchronously” with the match presents info about a soap opera, which will be presented hours later. This problem cannot be tackled using previous media, stream, or object layers, because it is a specification made on demand by the user in the moment that he sees the video, audio, and application. Similar cases happen with mash-up, interdestination, and context based applications.

As one of the many scenarios for multimedia applications, television (broadcast and broadband) has synchronization requirements in all these layers: intrastream synchronization to synchronize presentation of audio and video LDU unit streams; interstream synchronization to lip-sync audio and video; synthetic synchronization to provide interactive applications (Brazilian Ginga, European MHP... ) and enhanced content (subtitles...); and both IDMS and context synchronization for social sTV, second screen, and mash-ups applications. The focus of this study is to find approaches used to synchronize contents in TV scenario, characterizing the contents and the synchronizing solutions.

This paper is organized as follows: Section 2 introduces the methodology used to perform the systematic review; Section 3 presents the review's results; Section 4 shows a classification derived from the papers analysis; Section 5 briefly comments the research limitation; and Section 6 presents some conclusions.

## 2. Method

The need to synthesize available research evidence created well-established evidence-based disciplines [11] such as medicine and education research method called systematic literature review. This practice has recently been recognized in several computing disciplines from software engineering [12] to HCI [13]. More recently, a new method derived from systematic literature reviews was introduced: systematic mapping studies. Such studies are more focused on developing classification schemes of a specific topic of interests and reporting on frequencies of publications which cover a given topic of the development classification schemes.

This work reports the findings of a study that was conducted by combining the methods for systematic literature mapping and review to investigate the current state of

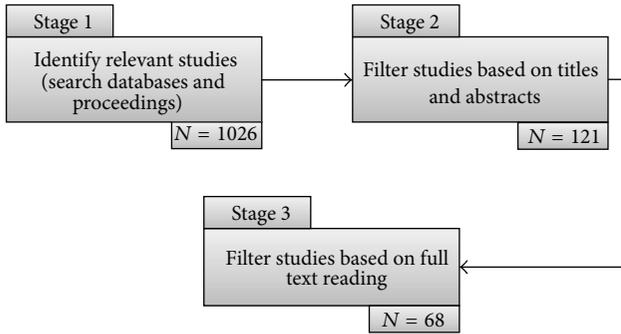


FIGURE 4: Study selection process and the number of works resulted from each stage adapted from [14].

research on Television Multiple Contents Synchronization. The details of research methods are described in the following subsections adapted from [14].

*2.1. Study Design.* This section presents the main focus and goals, points out the questions this review attempts to answer, and explains what research papers were included and excluded.

The focus of this literature review is based on Cooper's research outcomes, research methods, and practices or applications categories [15]. The research outcomes reveal gaps in the literature with regard to TV contents synchronization. The findings are based on the systematic analysis of data collection of research material. Research methods are analyzed to provide an overview of approach evaluations used by researchers and their contribution focus. The focus on practices and applications shows useful information regarding what type of content is provided in prototypes, where they came from, and where they are presented.

Our goal is to integrate outcomes and synthesize the results. We also attempt to generalize findings across the collected research papers. Another important goal of this review is to identify and characterize synchronization techniques for television multimedia environments and types of contents used in this synchronization, related to the semantic layer.

Finally, a set of important questions to be answered by this review is as follows.

- (i) What is the state of synchronization techniques for television multiple contents?
- (ii) Which devices are used to present multiple contents?
- (iii) What protocols and algorithms are used in the synchronization?
- (iv) What applications demand content synchronization?
- (v) What kind of synchronization is demanded by contents?
- (vi) What contents are being synchronized?

*2.2. Data Collection.* The data collection followed the process presented on Figure 4.

Stage 1 consists of a database search through academic and state-of-the-art publication databases and a manual

search in the proceedings of some of the main symposiums and congresses in TV and multimedia area. Four digital libraries were identified to be searched in a systematic manner:

- (i) Engineering Village (<http://www.engineeringvillage.com/>),
- (ii) ACM Portal Digital Library (<http://dl.acm.org/>),
- (iii) Scopus (<http://www.scopus.com/>),
- (iv) IEEE Xplore Digital Library (<http://ieeexplore.ieee.org/>).

They were chosen because own search engines allow the use of logical expressions or equivalent mechanism, include computer science publications or related topics that are related to the points being researched, allow the search within metadata of publications, and are accessible through the academic research network of the authors. These databases are commonly used sources for conducting systematic surveys in computing research. Not all databases had the same features and search capabilities, so it was necessary to apply modifications on the search for each specific library. The logical Boolean string used in the conducted search is listed below:

(tv OR television OR televisions)  
And  
(synchronisation OR synchronization OR synchronous)  
And  
(media OR multimedia OR stream OR flow OR  
content OR application OR applications)

To conduct the equivalent of the above search string it was necessary to learn how each of the digital library's advanced search features works. The end result was that all papers retrieved had within their title, abstract, or keywords the combination of the keywords presented in the logical Boolean string. Before proceeding to stage 2, all duplicated publications were removed. Figure 5 shows the intersections between the publications found in the databases and manual research. The number of papers selected after stage 1 was of 1026.

Exclusion and inclusion criteria were outlined in Stage 2 to filter irrelevant studies from Stage 1. The title and abstract of every paper were individually examined for false positives; that is, it was possible for a search result to contain all wanted keywords but without necessarily discussing the points of this review. At this stage, a total number of 121 studies remained. In this group works were included in which one had doubts about inclusion or exclusion. The exclusion and inclusion criteria were:

- Inclusion:
- (i) discusses solutions for synchronization contents in the television;
  - (ii) discusses cases that address the synchronization involving TV content;



TABLE 2: Publication sources of the final selected papers.

Publication source	Papers
International Conference on Consumer Electronics	7
International Conference on Multimedia and Expo	4
International Conference on Interactive Experiences for TV and Online Video	4
International Symposium on Consumer Electronics	4
Consumer Communications and Networking Conference	4
Brazilian Symposium on Multimedia and the Web	3
International Symposium on Multimedia	3
International Symposium on Broadband Multimedia Systems and Broadcasting	3
International Conference on Information and Communication Technology Convergence	2
Others	34

## (2) TV contents synchronization specific information:

- (a) transport of the synchronization specification;
- (b) synchronization channel;
- (c) synchronization mechanisms;
- (d) synchronization specification methods;
- (e) synchronization level;
- (f) sources and destination of contents;
- (g) control scheme;
- (h) qualitative and quantitative evaluation metrics;
- (i) applications and cases;
- (j) content characteristics;
- (k) paper's focus.

### 3. Results

This section presents the data extracted from publications resultant from stage 3. The general information is first presented about papers and then specific ones.

*3.1. General Overview.* The distribution of papers collected over the years is shown in Figure 7. It goes from 1998 with SMIL introduction [17] to the papers of 2013, like the work that extends the TV screen through projected screens on the wall around the TV [18].

As seen in Figure 7 most papers appear in the last five years, with a peak in 2012. In 2013 only two papers are seen up to half a year. The low number of papers can be explained by the fact that many proceedings of that year where most papers were found were not published to date.

Table 2 presents the main publications sources (Congresses, Symposiums and Journals) where papers were published.

*3.2. TV Contents Synchronization.* This subsection reports research results as determined by the analysis. Here aspects of synchronization, devices, sources, and contents are presented as they were extracted from papers.

The contributions of the surveyed papers are applications: the paper focus on the description of a specific multimedia application [18–30]; architecture: the paper proposes an architecture to solve synchronization problems but

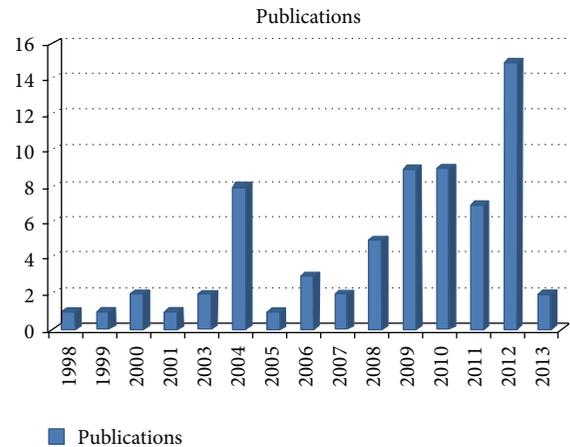


FIGURE 7: Publication years.

does not present programming interfaces or formalization [31–39]; framework: the paper presents a framework [40] that developers may use to provide synchronization to their multimedia applications [41–49]; language: the paper presents the description of a programming language that may be used to develop applications [17]; model: the paper presents the modeling of an approach that in theory may be used to bring synchronization to the applications [50–56]; platform/middleware: a platform which provides synchronization functionalities is presented. Also modifications to existing platforms are considered in this category [10, 57–69]; protocol: the paper defines rules and conventions for communication among devices so they may keep synchronization [70–74]; tool: the paper presents a tool with specific functionalities that once executed will provide synchronized contents [75–84]. The distribution of these papers is shown on Figure 8.

Platforms/middlewares are the majority kind of contribution found. The paper's authors propose a full environment to turn the presentation of contents with television possible. They contribute to the sources, transport, and presentation of contents and synchronization specifications. Hybridcast [67], HbbTV [69], and Ginga [10, 61] are some examples of middlewares. They are commercial platforms that are in

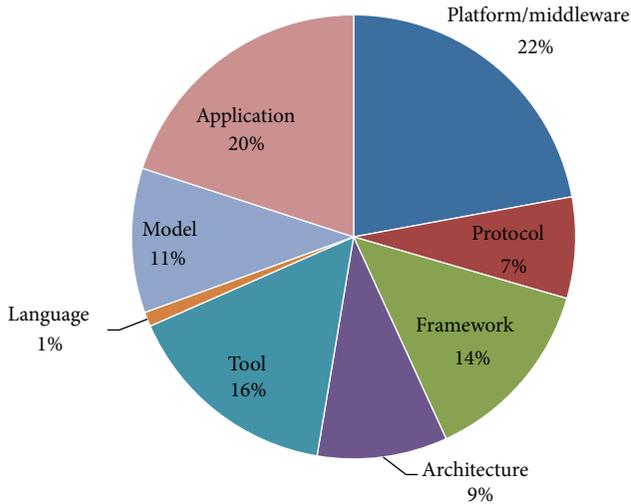


FIGURE 8: Focus of the contribution.

use in many countries: Hybridcast in Japan through NHK, HbbTV in European countries, and Ginga in Brazil and countries that adopted ISDB-Tb [10].

Application papers mainly present an experience around a specific multimedia application. Being the focus, details about implementation, interface, and tests are more detailed than other papers that only cite what kind of application was used to validate the proposal. There are many cases, but some can be highlighted: [19, 20] present solutions to show medical data within captured video; [24] presents an application to sign language education; [18, 27] present applications to enhance home entertainment exploring ubiquitous applications; and [25] presents the use of social networks application as a mean to measure the “heat” of a topic from TV.

Tools papers present applications which functionalities are not focused on the TV viewer but aim at other users, like the TV station. These papers themes are mainly focused in audience estimation [77, 84]; subtitles/closed caption/sign language automatic generation/synchronization [78, 80, 81, 83]; and video annotation [75, 76, 79].

**3.2.1. Synchronization.** When utilized, the term synchronization commonly means that something occurs at the same time as something else. This is confirmed by definitions of synchronization extracted from the selected papers.

Brunheroto et al. [31] present synchronization as loose or tight one. Loose synchronization typically depends upon the reception of a message (trigger) or the presence of data and does not require time stamps carried within the data encapsulation. Tightly synchronized data will require the presence of time stamps and careful control of emission and decode timing.

Park et al. [32] present synchronization as asynchronous, synchronous, and synchronized data. Asynchronous data has no time relation to main content being presented; on the other hand synchronous and synchronized data carries timing information so it can be linked to the main content. But

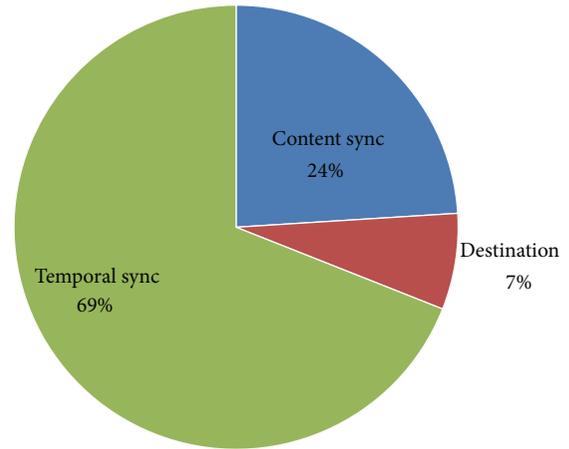


FIGURE 9: Synchronization focus.

the text does not present the difference between synchronous and synchronized.

Lai-Yeung and So [48] classify communication in synchronous versus asynchronous, synchronous: real-time communication among participants who are in different locations at the same time and asynchronous: communication over a period of time among participants in which the communication is characterized by a time lag among parties.

However in a more general sense some authors use synchronization as comprising content, spatial, and temporal relations between media objects [9]. This view presented by Blakowski and Steinmetz is one of the bases to classify synchronization in three categories (Figure 9): content, destination, and temporal synchronization.

Content synchronization papers [21–23, 25, 28, 42, 43, 45, 48, 58, 62, 75, 77, 82, 84] consider semantic relations among contents. These papers take in account the question of *what* is being presented in the main content and *how* to present/generate extra content. Papers consider what is being presented in main content to connect people [23, 25, 28, 42, 43, 48]; measure audience [77, 84]; connect other contents to the MC [21, 22, 45, 58, 75]; and personalize them [62].

Destination synchronization papers [10, 18, 26, 37, 52, 65] focus on *where* to present main content and related contents so they complement each other. Papers approach presentation on secondary devices [10], distribute the presentation of rich multimedia [37, 52], expand the TV screen in projected screens [18], or focus on distributed data for multiple devices [26, 65].

Most of papers approach the temporal synchronization, where the focus is to present all contents in a specific time interval, giving impression that they occur at the same time. The precision time of synchronization presented on papers was an objective of this review. The classification used to analyze the precision was based on [74]: very high synchronization (asynchronies lower than 10 ms), high synchronization (asynchronies between 10 and 100 ms), medium synchronization (asynchronies between 100 and 500 ms), and low synchronization (asynchronies between 500 and 2,000 ms). But as shown in Figure 10, only 15% of papers

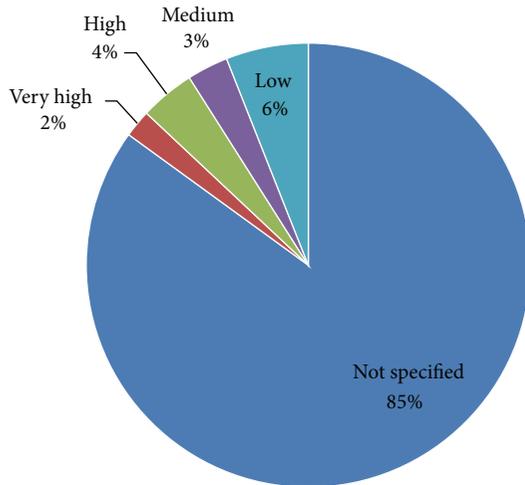


FIGURE 10: Synchronization precision.

presented enough data to achieve this goal; in the others the precision was not specified.

**3.2.2. Synchronization Specification.** The synchronization specification of a multimedia application describes all dependencies between its multimedia objects. Because the synchronization specification determines the whole presentation order and coordination, it is a central issue in multimedia systems [9].

Some important aspects related to synchronization specification are addressed next: transport of the synchronization specification, synchronization channel, synchronization specification methods, synchronization control scheme, and synchronization location.

**3.2.2.1. Transport of the Synchronization.** At destination, the presentation platform needs to have the synchronization specification at the moment that each object of the application is to be displayed. Three main approaches that support presentation synchronization are considered [9]: (i) pre-orchestration of the complete synchronization information before the start of the presentation, (ii) use of an additional synchronization channel, and (iii) use of multiplexed data streams.

Figure II shows the number of papers that used each approach. As seen most papers [20, 21, 27, 29, 31–35, 39, 41, 44–47, 49, 50, 54–56, 58, 59, 63, 64, 67–75, 78, 80, 81] have the synchronization specification multiplexed with the data streams. Sending it within the data stream implies that both media and specification are delivered together to the presentation device. The device can use this specification to play media synchronously. These papers commonly use MPEG based technologies (Subsection 3.2.3) or derivations to send the specification with the media.

The delivery of the complete synchronization information before the start of the presentation [18, 22, 24, 28, 30, 37, 48, 60, 61, 65, 66, 79, 83] implies that the full synchronization specification is delivered before any synchronous action

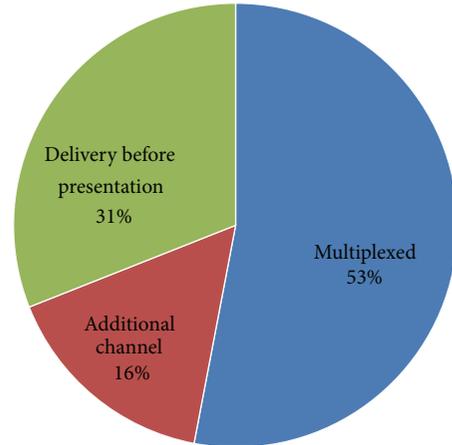


FIGURE 11: Transport of the synchronization specification.

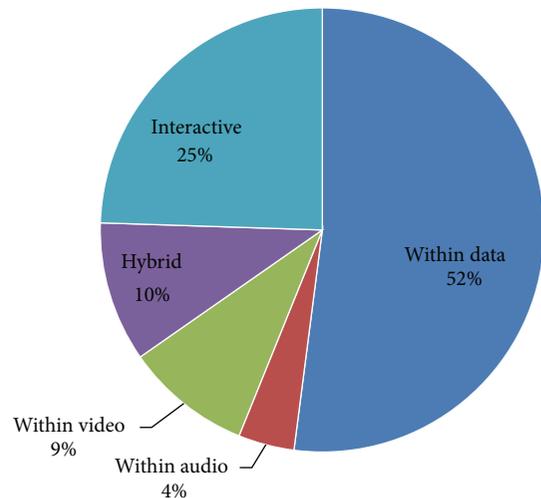


FIGURE 12: Synchronization channel.

is made. Examples of it are the use of NCL and SMIL languages to specify synchronization. In these cases the NCL and SMIL documents must arrive in the device before the synchronization rules starts.

Using an additional synchronization channel implies that the specification will arrive in a different channel than the one transmitting media [23, 25, 26, 36, 38, 42, 43, 62, 77, 82, 84].

**3.2.2.2. Synchronization Channel.** In this subsection papers are classified based on the channel used to transmit the synchronization specification. The possible channels are (i) an interactive one, within (ii) audio, (iii) data, or (iv) video, and (v) a hybrid approach (Figure 12).

In the case of interactive channel, the multimedia application uses its capabilities of communication with different servers to retrieve media and the specification required to present media synchronously with a main content [22–26, 36, 37, 42, 45, 48, 51, 52, 62, 64, 77, 79, 84]. Some papers use this channel as a way to receive the specification only [26, 64], but

in most cases, besides the specification, the other media are received through this channel.

In a broadcasting transmission, three possibilities to send sync information to the presentation device arise: within audio, video, or data. ASR (Automatic Speech Recognition) can be used on the audio to retrieve the speech of the TV show and this info can be used to synchronize and generate extra content [80]. Another alternative is to use a sample of audios as a dynamic anchor to achieve synchronization among contents [38, 43]. At last anchors can be sent multiplexed with the main content's audio in a way that users does not note the modified audio, but applications may listen to this audio and use it to synchronize contents (audio watermarking) [64].

Within video [20, 29, 30, 41, 70, 82], the synchronization specification can be sent directly in the video, where both user and applications can see the anchors for synchronization as presented in [82], which uses QRcodes to make the synchronization and in [30] there is a call inviting the user to connect his phone with the application. [29] presents a solution using steganography where only the application notices the anchors used in the synchronization. [20, 70] personalize mpeg standards introducing information into the video frames and extracting it before presentation. [41] uses digital image processing to track video objects and create multimedia anchors.

Within data [10, 17–19, 21, 27, 28, 31–35, 39, 44, 46, 50, 53–56, 58–61, 66, 68, 71–76, 78, 81, 83], the specification is included with the data sent with the broadcasting. In this case the specification is sent with the protocol headings (e.g., RTP/RTCP in [73]), as content transferred by the transmission (e.g., Ginga documents [10]) or as metadata information (e.g., used with mpeg solutions [46]).

In hybrid approaches the synchronization specification is sent through both broadcasting and an alternative channel. [63, 65] present solutions that use the second screen concept (the use of a second device, besides TV, to interact with the content presented in TV screen) to synchronize contents. In this case the second screen communicates with TV to receive the synchronization specification from it and also communicates with a remote server to receive the synchronization specification for the extra content. [49, 67, 69] use the hybrid platforms (HbbTV and Hybridcast) to synchronize contents. These platforms directly receive the broadcasting and broadband contents and synchronize them with its specifications. MITv [57] is a platform that sends interactive content within the broadcasting or an interactive channel. The channel used depends on the demand of the extra content: if the demand is huge it uses the broadcasting, if not it uses the interactive channel. Margalef et al. [47] proposes an interactive platform for DVB-H that sends interactive content through interactive channel and receives main content through broadcasting.

**3.2.2.3. Method.** For the specification of multiple object synchronization, including user interaction, various specification methods must be used [9]. Table 3 shows how the selected papers were categorized among the six synchronization methods.

In event-based synchronization, the presentation actions are related to synchronization events [26, 30, 36, 45, 53–55, 66,

TABLE 3: Synchronization specification methods.

Method	Number of Papers
Event-based synchronization	9
Hierarchical specification	16
Reference points	30
Synchronization based on a global timer	10
Synchronization based on virtual axes	4
Contextual rules	2

77]. Typical presentation actions are start and stop a media presentation, wait a user interaction, and so forth.

In hierarchical synchronization, media objects are regarded as a tree of nodes [10, 17, 22, 28, 37, 42, 48, 51, 52, 58, 60, 61, 76, 79, 83]. Hereinto, the leaf node can be single media processing and also can be user input or delay. Hierarchical structure is easy to compute storage and handle, so it has been widely used. The limitation of hierarchical structure is that each movement only can be synchronized in its beginning and end [85].

For synchronization based on a global timer, all objects are attached to a time axis that represents an abstraction of real time [19, 20, 27, 35, 49, 69, 71–73]. In virtual time axes specification method, it is possible to specify coordinate systems with user-defined measurement units [56, 59, 63, 64].

In the case of synchronization via reference points, objects are regarded as sequences of LDU's [18, 21, 24, 29, 31–34, 38, 39, 41, 43, 44, 46, 47, 50, 57, 67, 68, 70, 74, 75, 78, 80–82, 84]. The start and stop times of the object are called reference points.

In [23, 65] synchronization is dictated by the use of contextual relations. When a specified contextual situation occurs, like the use of users position [23], a synchronization action takes place.

**3.2.2.4. Control Scheme.** Generally, three schemes are employed to perform synchronization control (Figure 13) [74]: two centralized schemes (i) Master/Slave or M/S Scheme and (ii) Synchronization Maestro Scheme or SMS and (iii) one distributed scheme, Distributed Control Scheme or DCS. Besides the three control schemes, this review adds two derivations for the SMS Scheme that are described next: (i) Blind Maestro and (ii) Passive Producer. Papers were classified using these schemes considering broadcaster's main content as media source and user devices as receivers (Figure 14).

Works that use DCS [36, 59, 64, 65, 79] have all the receivers multicast feedback information about their playout to all the other receivers and each one of them selects the synchronization reference from among its own playout timing and those of the other receivers.

In M/S Scheme [19, 20, 22, 27, 32, 42, 56, 72, 73], receivers are differentiated into master and slave. The master receiver multicasts feedback control messages about playout timing to all the slave receivers. Accordingly, each slave receiver adjusts its own playout process to the reference playout process of the master.

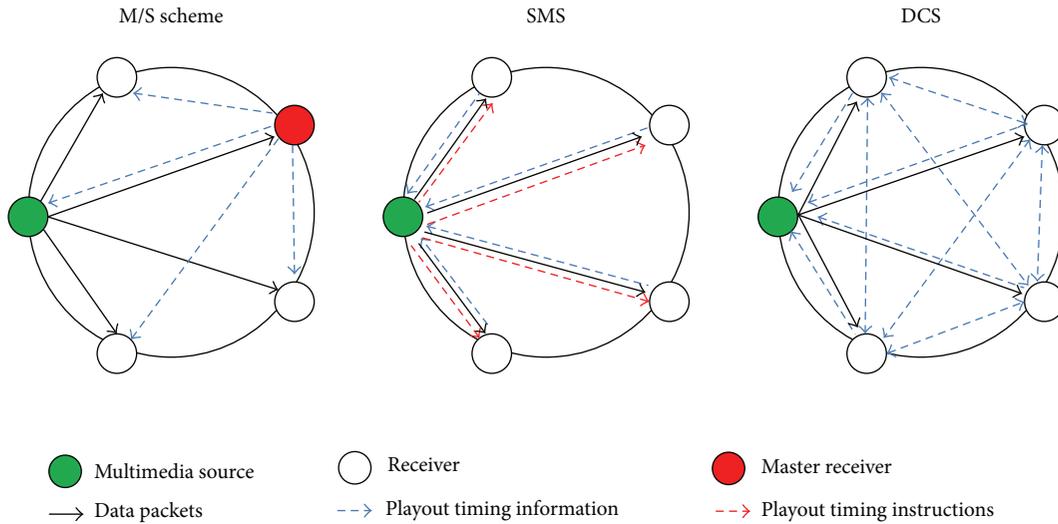


FIGURE 13: IDMS control schemes, [74].

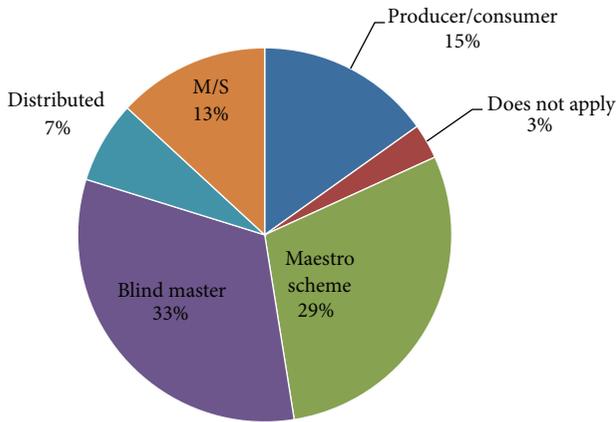


FIGURE 14: Synchronization control scheme.

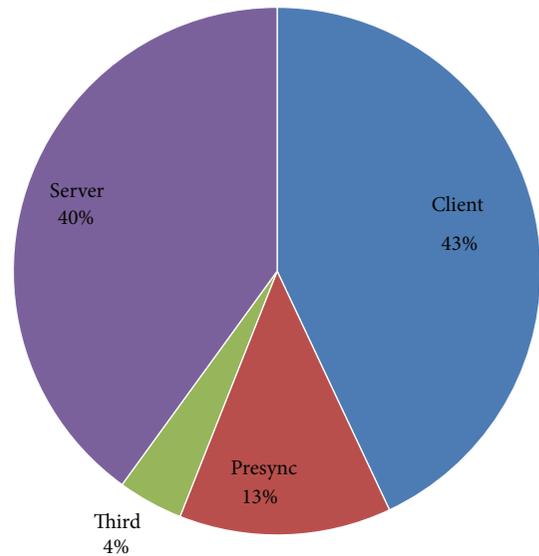


FIGURE 15: Location of synchronization.

Papers with SMS [10, 17, 18, 23, 24, 33, 37, 47, 49, 51, 52, 54, 61–63, 67–69, 71, 74] uses the existence of a synchronization maestro or manager (that can be the source, one real or fictitious receiver, or a completely separate entity), which gathers the playout information of the receivers and corrects their playout timing by distributing new adapted control messages. Set-top-boxes and TVs commonly play the role of the master once they directly receive the broadcasting content and can communicate with the other devices.

The Blind Maestro Scheme [28–31, 34, 35, 39, 41, 44–46, 50, 53, 55, 57, 58, 60, 66, 70, 76, 81, 83] differentiates from SMS because the Blind Maestro sends the synchronization specification to all devices connected to the broadcasting channel, being millions of devices or none. In other words, the maestro does not know who he is coordinating, but regardless of this information he keeps sending all specification he can to everyone. This can be a common scenario in DTV systems, where the broadcaster sends main content and extra content to receivers without knowing who will receive the signal.

In the case of the Passive Producer Scheme [21, 25, 38, 43, 48, 75, 77, 78, 82, 84] the content provider does not send direct synchronization specification to the receivers within the broadcasted content, but this content is used by another entity to generate the synchronization specification. It is passive because it does not generate synchronization points on purpose but produces the content that is used to generate the specification. An example is the use of audio fingerprinting techniques. Anyone can receive the main content (audio and video) and generate the fingerprinting without the direct intervention of the main content producer.

*3.2.2.5. Location of Synchronization.* The synchronization of multiple contents can happen at four different places: (i) at the server, (ii) at the client, (iii) in an external entity (third), or (iv) presync on server (Figure 15).

Synchronization on server side [17, 19–21, 28, 31, 32, 34, 41, 42, 44, 46, 48, 50, 55, 57, 58, 60, 62, 66, 70, 75, 76, 78, 80, 81, 83] maximizes client's bandwidth, because only one multimedia stream is delivered to client [35]. All extra content is sent synchronized with the main content and just played on the client.

Client-side synchronization [10, 18, 22–27, 29, 30, 33, 35–37, 43, 49, 52–54, 56, 59, 61, 65, 72–74, 77, 79, 82] requires more client bandwidth (two or more multimedia streams delivered to client) but presents more options to personalize extra content by the user [35]. The synchronization is performed using the client devices and information available at his side.

A third alternative is to presynchronize the contents on server and then resynchronize then on the client [39, 45, 47, 51, 63, 64, 67–69]. In this case there is a synchronization of contents both on server and client side. This happens because the server gathers all contents but transmits them in separated channels that will converge again on the client side.

The third synchronization presents an external entity besides server and client that is responsible for the synchronization of contents to be presented on the client. Stokking et al. [71] proposed a Media Synchronization Application Server (MSAS) that is responsible for the synchronization of different clients; it collects synchronization status information from them, calculates delay settings instructions, and sends these instructions to the clients. In [38, 84] samples of the main content are sent to a third synchronizer that uses this sample to measure audience and provide second screen synchronous applications.

**3.2.3. Technologies.** The survey explored the main technologies (middlewarees, protocols, platforms ...) applied in the sixty-eight selected works. Among all technologies the following are highlighted: MPEG standards, Real Time-Transport Protocol (RTP), Network Time Protocol (NTP), ISDB-Tb (Ginga), Multimedia Home Platform (MHP), Hybrid Broadcast Broadband TV (HbbTV), and QR Code.

Thirty-four percent of works (twenty-three papers) present solutions directly associated with one of Moving Pictures Experts Group defined standards. The most adopted [32, 33, 35, 41, 44, 50, 51, 63, 64, 66, 67, 69] standard is the MPEG-2, that consists of the standard for the generic coding of moving pictures and associated audio information. MPEG-4 that defines digital compression of audio and video is part of the solution for five papers [34, 41, 46, 47, 56] that adds modification for the standard or uses it as format for streaming distribution. MPEG-7 is the standard for multimedia content description used in the works [23, 41, 75] to describe the main content and with it correlate extra content. MPEG-21 is a suite of standard that defines a normative open framework for end-to-end multimedia creation that aims to benefit the content consumers providing them access to a large variety of content in an interoperable manner. It is used by [28, 52, 59] to generate contents that may be used in multiple devices and situations. Finally, MPEG-DASH provides a solution for the efficient and easy streaming of multimedia using existing available HTTP infrastructure. It is used in [68] as

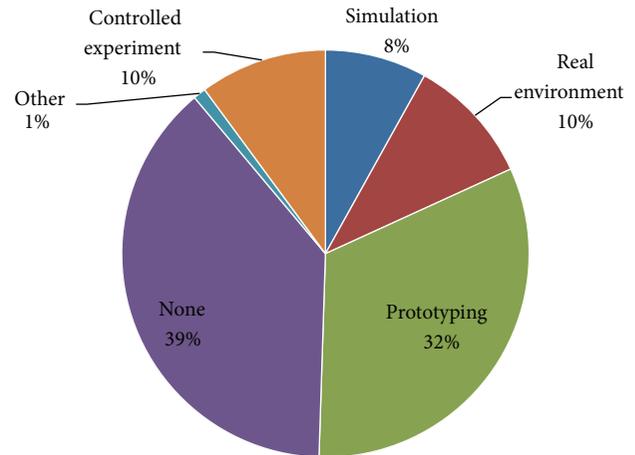


FIGURE 16: Evaluation metrics.

opportunity to maintain compatibility to HbbTV platform specifications.

RTP defines a standardized packet format for delivering audio and video over IP networks. Fourteen papers [34, 35, 46, 47, 49, 55, 56, 59, 64, 69, 71–74] use the advantages presented by RTP [86] in two situations: for distribution of the main content among different clients or delivery of audio and distribution of extra contents. With the use of RTP the NTP is commonly used (a networking protocol for clock synchronization between computer systems over data networks). Among the fourteen papers that use RTP as part of the solution, half [27, 35, 49, 71–74] also use NTP to permit a global clock synchronization for all involved entities.

Three platforms that provide the possibilities of extra contents presentation within main content were found: Ginga [24, 28, 29, 53, 55, 60, 61, 76, 83], MHP [41, 44, 51, 59], HbbTV [49, 68, 69], and HybridCast [67]. Both Ginga and MHP are middlewarees used mainly for interactivity and presentation of contents in multiple devices. HbbTV works focus on the convergence of broadband and broadcasting contents on the user's television and HybridCast provide television (TV) programs with rich and varied applications.

**3.2.4. Evaluation.** This subsection presents the evaluation methods used or not in the surveyed works and the cases used as base. The evaluations were done by use of (Figure 16): controlled experiments; prototyping; real environment tests; simulation or none if no evaluation was done (twenty five papers).

Controlled experiments are investigations with groups that study variables that may affect or influence one or more factors of the proposed work. Seven were identified [18, 30, 36, 38, 76, 79, 83]. Most of them [18, 30, 76, 79, 83] focus on the investigation of the final result, which means the evaluation of the application functionalities and interface aspects and not on the synchronization aspects that may affect the results. But the works [36, 38] focus on the evaluation of the synchronization itself.

Vaishnavi et al. [36] presented two experiments that focus on identification of the skew tolerance for different content

presentation, in the case of a social TV application for real time voice and text communication. In the experiments users watched a video together at two different locations, communicating with each other using text chat or voice chats. At specific intervals, the synchronization of the videos was automatically changed in a way that one of the participants was not in sync with the others and the impact of this change was measured. Duong et al. [38] presented an experiment in order to validate the performance of the proposed system. They made controlled recordings in living rooms varying distance and noise of recordings. They evaluated synchronization performance in terms of precision (the fraction of detected synchronization positions that are correct) and recall (the fraction of correct synchronization positions that are detected).

Prototyping means to produce a first or experimental working model of something proposed. Twenty-two works [10, 23, 24, 28, 29, 33, 34, 42, 44, 46, 47, 51, 58, 60, 61, 63, 64, 66, 69, 70, 81, 82] used this approach to validate their proposal. The use of a real environment as an evaluation technique is a step further of prototyping. In this case a final tool or model is produced and placed to be used by real users. Platforms [67], TV station extra content generation [21, 78] and audience measurement [77] tools, social networks data input [25], and main content analysis [43] are examples of works that use solution in real environment tests.

Instead of using prototypes or a real environment, researches can use simulation: a representation of the problem, attempting to predict aspects of the behavior of the system by creating a model and simulating it on a virtual environment. In TV scenario, researches [57, 59, 68, 75, 80, 84] simulate the broadcasting transmission with a local video file or stream server and add extra content with the use of other local server.

For the evaluation, different cases were used as example for future application of the proposals. Entertainment is presented as study case in twenty-five works, twelve about sports (additional audio streams, automatic annotations, games statistics, betting platforms, and others) [21, 22, 42, 49, 53, 58, 66, 69, 71, 75, 76] and thirteen about general entertainment (movies, series, soup-operas, etc.) [10, 18, 29, 33, 34, 43, 47, 48, 57, 60, 61, 70].

Accessibility applications help people with disabilities or reduced capabilities to watch and interact with TV contents. This kind of application is presented with two different focuses: closed caption and subtitles and descriptive audio and sign language. [78, 80, 81] present techniques that generate automatic subtitling or closed caption related to the main content, synchronizing both before the transmission. On the other hand, [67] presents a solution for showing subtitles and others contents on client side from multiple sources. Each paper that approached sign languages presented a different scenario: [63] proposes that extra content (sign language) and main content are sent through different channels and that they are synchronized at client; in [24] the synchronization among contents relies on a Web environment that is used to generate synchronization point between main content and the sign language signal; and in [83] main content is sent embedded with reference points that are used to play

synchronously sign language videos related to these reference points that are stored locally on client side.

Social TV is a label for Interactive TV (iTV) systems that support the sociable aspects of TV viewing [87]. [23, 25, 36, 43, 67, 79] explore this aspect and promotes to users possibilities to: share contents, chat via text and voice, create ad hoc communities based on TV watching, and annotate videos for temporal asynchronous watching.

In personalization cases on-demand personalized content is rendered on user device (TV, smartphone, and others) synchronized with content on a receiver [28, 64]. Ubiquitous home includes synchronization not only of audio, video, text but also synchronizing other information with user peripheral devices [54, 56].

Other cases include t-health [19, 20, 43], news [51], t-commerce [33, 44, 60, 62] an educational applications [30, 46, 82], or no study cases [17, 26, 27, 31, 32, 35, 37–39, 41, 45, 50, 52, 55, 59, 65, 68, 72–74].

*3.2.5. Content.* Contents are information that may provide extra value for an end user in specific contexts. They may be generated by single or multiple sources and presented on one or many devices. They may be generated in real-time or offline. Real-time contents [31, 45] are information that is generated at the same time that the main content is. Offline [31] or other [45] contents are information that are generated before the presentation of the main content.

Next, main sources and presentation devices used to present and generate contents found in selected papers are described.

*3.2.5.1. Source.* A source is the point or place from which something is originated. In this case, source is the place where contents originates and are sent to TV viewers.

While watching TV the viewer may have access to multiple contents besides the main content presented on TV. This mapping shows that 88% of papers use multiple sources as input to TV viewer experience. Among the sixty-eight papers, only one does not use the main content produced by a broadcaster as an input, in Fawcett et al. [22]. This latest paper do not use the main content as input because it is considered that the viewer is present on the same place where the program is happening, in the specific case, the viewer is in the football stadium watching the match while receiving the extra content in his device.

Other sources used are extra contents sent within the main content [10, 17, 28–32, 34, 41, 44, 46, 48, 53, 55, 57, 60, 61, 64, 70, 83]; extra contents provided by the broadcaster's web servers (interactive channel) [24, 33, 39, 45, 47, 58, 63, 66, 67]; extra content provided by web servers indicated by the broadcaster [18, 35, 37, 43, 51, 58, 65, 69]; independent extra content providers [39, 62, 64, 67, 68, 82]; and extra content generate in the client side: contextual information [19, 20, 38, 54, 65, 77, 84] and user generated information [23–25, 36, 42, 76, 79]. Contextual information uses information related to the user and collected from his environment, such as geolocalization, and social networks profiles. User generated



FIGURE 17: TV sharing [23].

content are information that he generates to interact with content and other users, like audio or textual messages.

Considering relation between extra and main content almost all works presented related multiple contents (97%), only two papers did not: Cheng [37] uses TV only as one of many possibilities of presentation of contents; he does not consider the MC in its work. In [65] the Maps-TV application for digital TV ignores the main content being presented on the device to present a collaborative map. In this case the multiple contents come from the application, user's contribution, and web services that feed the map.

**3.2.5.2. Presentation.** The contents sent from sources to the viewer can be presented in different devices. Cheng [37] describes three practical use cases of the use of multiple presentation devices:

- (i) device shifting: a single user would have the same content over different devices at different times;
- (ii) companion device: user can control multiple devices at the same time;
- (iii) device collaboration: users expect to share their interactions and content with others on multiple devices in a collaborative way.

By targeting small and multiple devices, IDTV content can become portable, for “anytime, anywhere” interaction [33]. This concept is important in situations like the one shown in Figure 17, where multiple user have shared main screen but each one can have a personal interactive device.

Many devices are presented as possible presentation devices for the multiple (main and extra) contents provided by the sources. The devices used in the surveyed papers are television, personal computer, smartphones, PDAs and Tablets. The main content is commonly presented on TV, and when it is not the case, the PC is used to simulated the TV and add functionalities. The other devices are used to present

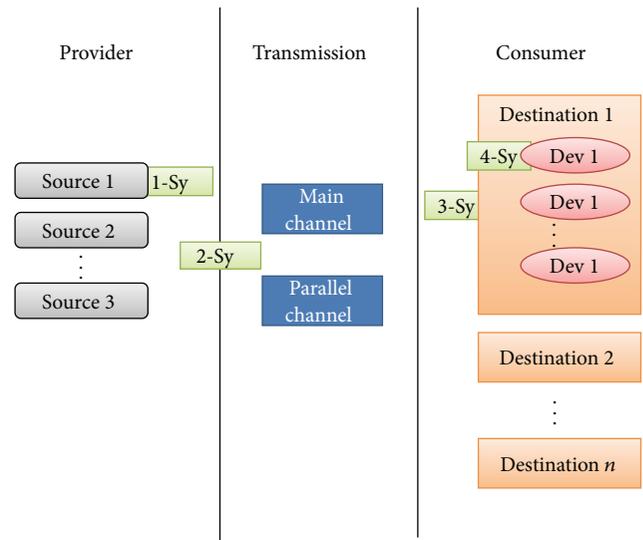


FIGURE 18: Overall scenario.

extra contents, like additional audio tracks, web contents, text chats, and others that were previously presented in this text.

## 4. Overview

This section discusses subjects related to the analysis of the mapping.

**4.1. Scenario.** Multiple sources, several destinations, different devices, and diversified synchronization points. These points show how complex the definition of a synchronization scenario can be. An overall view of the scenarios found in this survey is presented Figure 18.

Figure 18 presents three lanes: the provider, the transmission, and the consumer.

The provider represents the content providers. Each content provider is a source of contents that are available to consumers through the transmission. Each source is responsible for one type of content (e.g., the main content, subtitles, sign language, web related information, etc.). Primary sources are the ones that generate their own contents and transmit them. Secondary sources unite contents from different primary sources and transmit them to consumers.

The transmission lane represents the transmission means of contents. Two possibilities are presented: main channel and parallel channel. The main channel carries the main content and is broadcasted to all consumers. In other words, all contents sent with the broadcasting can be accessed by any consumer. With the main content, some extra content can be sent through the main channel, like subtitles, alternative audio, and others. The parallel channel on the other hand has a communication link with each consumer that desires to receive content through that channel. This channel is used as an alternative route to send extra contents from sources to clients.

The consumer lane presents two subjects: destination and device. A destination defines a viewer and its environment

that may be composed of one or several devices (TV, smart-phone, PC, etc.) that will enhance user experience helping to retrieve, play, and interact with contents.

Figure 18 also presents four synchronization points (Sy—green box). These synchronization points may be located on different places and are classified as:

- (i) master source: one of the sources accesses the others and synchronizes them with its own content;
- (ii) content processor: an entity that is not a source gathers different sources and generates synchronization points among them, so clients can use this information to present the different contents synchronized;
- (iii) client gateway: at consumer side, a gateway receives all contents and synchronizes them based on their specification. The gateway then distributes the contents to the different applications and devices;
- (iv) master device: one device is able to synchronize the contents that it is playing with the contents presented on other devices. If it can communicate with other devices it may work and also coordinate them.

**4.2. Classification Scheme.** Based on the mapping, this section synthesizes a classification scheme that shall be used in future complementary surveys and to classify related papers in the area. The scheme is constituted by the following.

**Synchronization Type.** It describes which synchronization relation is approached in the paper: time, destination, or context. Time relations consider that multiple contents must be presented in a limited time interval to be synchronized. Destination relations consider the relationship of contents over multiple devices, where each content should be presented, and how to migrate contents over different devices and the collaboration among them. Contextual relations consider the semantic relations among contents, if the contents being presented are correlated, how to offer an extra content related to the main content, how to use viewers information to enhance their experience, and so forth.

**Synchronization Specification.** Synchronization specification of a multimedia object describes all dependencies and rules related to the media presentation. Some information may be extracted from paper to make it easier to understand their specifications.

- (a) Transport: at destination, the presentation components need to know the synchronization specification at the moment a media is to be displayed. This specification is delivered to the presentation players: delivered before the start of the presentation, using an additional synchronization channel or multiplexed in data streams.
- (b) Channel: the channel defines what means of transmission was used to send synchronization specification from the entity responsible for the synchronization to the ones responsible for playing them. They are

sent within the main content's data, audio, or video, through a parallel channel, or both.

- (c) Method: it describes the specification methods used to synchronize the multiple contents.
- (d) Control Scheme: synchronization techniques may differ in the form of controlling the synchronization. The players may distribute information to maintain synchronization or centralize the control on one player or entity. Section 3.2.2.4 presents the schemes related to iTV scenario.
- (e) Location: the synchronization of multiple contents can happen at four different places: at the media server, at the client, with an external entity (third), and presync on server. Different locations imply different responsible entities and consequently requisites that differ for each approach.

**Content information:** as these works focus on synchronization of multiple contents on the same or multiple destinations, describing these contents and their destination helps to characterize them.

- (a) Source/Media: sources may vary from a TV broadcaster to an user context information like his geolocation. Each source will present its own characteristics: a TV station will probably broadcast its contents, using different means like air, cable, or even IPTV, targeting anyone that symphonizes the television on their channel. If the source is one person geolocation, this information (content) will probably be used only to interact locally, with persons that share same location or other sources may use this information to personalize or analyze the person behaviors.
- (b) Destination: the destination, consumers, viewers, or clients are on the opposite side of the sources. They consume the contents generated by sources presenting them to the users. Just like sources, destinations in different environments can be composed by different devices. Each device has its own characteristics, as examples: an analogical TV can present only analogical audio and video; a digital TV can present digital audio, video, and data (EPG and interactive applications); smartphones may not play broadcasted audio and video but can play web streams and application. This environment will characterize how the user interacts and view the TV programs and also define its limitations (an additional audio track may be being offered in a RTP stream, but the user has no device able to play such stream).

## 5. Limitations of the Review

The limitations presented by this mapping are similar across other systematic reviews as well. There is likely some important material that was not included in the review such as dissertations, related books or white papers and some relevant papers might not have been found in the digital databases using the search and selection protocol.

## 6. Conclusion

This work aimed to present a systematic review about multiple contents synchronization for television scenarios. As result it presented the state of synchronization techniques for multiple contents presentation on television, showing papers related to the theme and their characteristics. The content characteristics shows what types of contents are used in the researches, where they are presented, what is their origin, how synchronization works, and the multimedia applications. Considering synchronization aspects, this work revealed synchronization related to synchronization specification issues and the aspects of synchronization based on the three relations: time, destination, and context.

Through the paper consolidated solutions like sending all content multiplex in the main channel are presented, but as new challenges arises, these solutions become limited and new ones are proposed. One example is the introduction of second screens. The second screen crashes all multiplexed contents solutions because other devices are involved in the process, not only tv, and these devices are able to retrieve its own contents. With this new device on TV scenario, sending all contents in main channel is not possible; once with second screen the user will see and interact with the content in a different device than TV, which many times has no direct communication with the second screen. Also the second screen offers to the user the possibility of personalized contents, something not possible if all contents are sent in the same channel. To solve this and other problems, solutions using hybrid approaches and interactive channel are presented.

The TV is no more a single device that plays content to user; it is now one device in a new environment constituted by multiple devices that are used to play diverse contents to users from different and independent sources. Synchronizing these devices and their contents are the challenge to be solved; once as seen the focus of most papers is to synchronize contents only in time. As shown in introduction, problems around context synchronization are open and the use of multiple devices is a hot topic around opening space for the multiple destination synchronization.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgment

The authors thank FAPESB (Fundação de Amparo à Pesquisa do Estado da Bahia) for financial support for the first author.

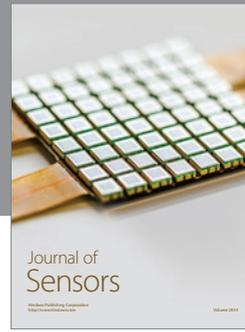
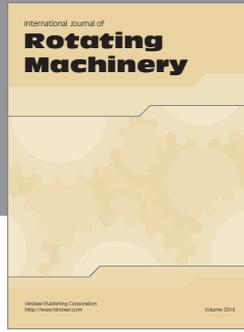
## References

- [1] P. Cesar and D. Geerts, "Past, present, and future of social TV: a categorization," in *Proceedings of the IEEE Consumer Communications and Networking Conference (CCNC '11)*, pp. 347–351, January 2011.
- [2] R. Malhotra, "Hybrid broadcast broadband TV: the way forward for connected TVs," *IEEE Consumer Electronics Magazine*, vol. 2, no. 3, pp. 10–16, 2013.
- [3] A. Baba, K. Matsumura, S. Mitsuya et al., "Seamless, synchronous, and supportive: welcome to hybridcast: an advanced hybrid broadcast and broadband system," *IEEE Consumer Electronics Magazine*, vol. 1, no. 2, pp. 43–52, 2012.
- [4] P. Cesar and K. Chorianopoulos, "The evolution of TV systems, content, and users toward interactivity," *Foundations and Trends in Human-Computer Interaction*, vol. 2, no. 4, pp. 279–373, 2008.
- [5] T. Meyer, W. Effelsberg, and R. Steinmetz, "A taxonomy on multimedia synchronization," in *Proceedings of the 4th Workshop on Future Trends of Distributed Computing Systems*, pp. 97–103, 1993.
- [6] H.-P. Dommel and S. K. Verma, "Multipoint synchronization protocol," in *Proceedings of the IEEE International Conference on Systems, Man and Cybernetics (SMC '04)*, vol. 5, pp. 4631–4636, October 2004.
- [7] Y. Ishibashi, A. Tsuji, and S. Tasaka, "Group synchronization mechanism for stored media in multicast communications," in *Proceedings of the 16th IEEE Annual Conference on Computer Communications (INFOCOM '97)*, vol. 2, pp. 692–700, April 1997.
- [8] R. Steinmetz, "Multimedia synchronization techniques: experiences based on different system structures," in *Proceedings of the 4th IEEE ComSoc International Workshop on Multimedia Communications*, pp. 306–314, 1992.
- [9] G. Blakowski and R. Steinmetz, "A media synchronization survey: reference model, specification, and case studies," *IEEE Journal on Selected Areas in Communications*, vol. 14, no. 1, pp. 5–35, 1996.
- [10] L. F. G. Soares, R. M. R. Costa, M. F. Moreno, and M. F. Moreno, "Multiple exhibition devices in DTV systems," in *Proceedings of the 17th ACM International Conference on Multimedia (MM '09)*, pp. 281–289, ACM, October 2009.
- [11] S. Straus, *Evidence-Based Medicine: How to Practice and Teach EBM*, Elsevier/Churchill Livingstone, 2005.
- [12] T. Dybå, T. Dingsøy, and G. K. Hanssen, "Applying systematic reviews to diverse study types: an experience report," in *Proceedings of the 1st International Symposium on Empirical Software Engineering and Measurement (ESEM '07)*, pp. 225–234, September 2007.
- [13] M. N. Islam, "A systematic literature review of semiotics perception in user interfaces," *Journal of Systems and Information Technology*, vol. 15, no. 1, pp. 45–77, 2013.
- [14] G. Hachey and D. Ga?evi?, "Semantic web user interfaces: a systematic mapping study and review," Submitted to *Semantic Web—Interoperability, Usability, Applicability an IOS Press Journal*.
- [15] H. M. Cooper, "Organizing knowledge syntheses: a taxonomy of literature reviews," *Knowledge in Society*, vol. 1, no. 1, pp. 104–126, 1988.
- [16] F. Boronat, J. Lloret, and M. García, "Multimedia group and inter-stream synchronization techniques: a comparative study," *Information Systems*, vol. 34, no. 1, pp. 108–131, 2009.
- [17] P. Hoschka, "An introduction to the synchronized multimedia integration language," *IEEE Multimedia*, vol. 5, no. 4, pp. 84–88, 1998.
- [18] R. Vatavu, "There's a world outside your TV: exploring interactions beyond the physical TV screen," in *Proceedings of the 11th european conference on Interactive TV and video (EuroITV '13)*, pp. 143–152, 2013.

- [19] S. B. Kim, Y. H. Lee, J. H. Kim, and S. I. Kim, "A full digital video eeg system," *Engineering in Medicine and Biology*, vol. 1, 1999.
- [20] S. Zeng, J. R. Powers, and H. Hsiao, "A new video-synchronized multichannel biomedical data acquisition system," *IEEE Transactions on Biomedical Engineering*, vol. 47, no. 3, pp. 412–419, 2000.
- [21] N. Nitta, N. Babaguchi, and T. Kitahashi, "Automated annotation of significant scenes based on structure of broadcast sports video," *Electronics and Communications in Japan*, vol. 86, no. 12, pp. 69–78, 2003.
- [22] J. Fawcett, B. Beyer, D. Hum, A. Ault, J. Krogmeier, and C. Taskiran, "Rich immersive sports experience: a hybrid multimedia system for content consumption," in *Proceedings of the 6th IEEE Consumer Communications and Networking Conference (CCNC '09)*, pp. 1–5, January 2009.
- [23] E. Marilly, A. Germaneau, B. Boidart, and O. Martinot, "Active TV: enter into a cross media, community-based experience," in *Proceedings of the 6th IEEE Consumer Communications and Networking Conference (CCNC '09)*, January 2009.
- [24] C. S. Oliveira, E. B. Dos Santos Jr., and H. C. de Oliveira, "Synchronizing Brazilian Sign Language (LIBRAS) on videos for open digital TV with educational purposes," in *Proceedings of the European Conference for the Applied Mathematics and Informatics*, pp. 18–23, December 2010.
- [25] D. Y. Wohn and E.-K. Na, "Tweeting about TV: sharing television viewing experiences via social media message streams," *First Monday*, vol. 16, no. 3, 2011.
- [26] L. B. Chen and M. Prokopi, "Resource-aware pairing device framework for ubiquitous cloud applications," in *Proceedings of the International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, pp. 252–258, 2012.
- [27] C. Cho, S. S. Kim, and K. D. Moon, "Enhancing home entertainment experience with virtual UPnP media renderers," in *Proceedings of the International Conference on Information and Communication Technology Convergence*, pp. 270–271, 2012.
- [28] I. Norões, V. Furtado, and E. Furtado, "Maps-TV: a software architecture to support development of collaborative maps in DTVi," in *Proceedings of the 10th European conference on Interactive tv and video (EuroITV '12)*, pp. 165–173, 2012.
- [29] M. Maritan, "Um Estudo sobre o Uso de Marcações Auto-contidas em Vídeo para Promoção de Interação Sensível ao Conteúdo em Múltiplos Contextos," in *Proceedings of the 19th Brazilian Symposium on Multimedia and the Web (WebMedia '13)*, 2013.
- [30] D. S. Martins, L. S. Oliveira, and M. G. C. Da Pimentel, "Designing the user experience in iTV-based interactive learning objects," in *Berlin 28th ACM International Conference on Design of Communication (SIGDOC '10)*, pp. 243–250, September 2010.
- [31] J. Brunheroto, R. Chernock, P. Dettori et al., "Issues in data embedding and synchronization for digital television," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '00)*, vol. 3, pp. 1233–1236, August 2000.
- [32] S. M. Park, Y. J. Lee, J. H. Choi, and J. S. Choi, "The design and implementation of data server for data broadcast service," in *Proceedings of the 10th IEEE International Conference on Electronics, Circuits and Systems (ICECS '03)*, pp. 1176–1179, December 2003.
- [33] M. Ma, D. Wilkes-Gibbs, and A. Kaplan, "IDTV broadcast applications for a handheld device," in *Proceedings of the IEEE International Conference on Communications*, vol. 1, pp. 85–89, June 2004.
- [34] L. Zhang, J.-Z. Cui, and J.-X. Zhang, "A method for storage and transport of embedded rich media application," in *Proceedings of the 1st International Symposium on Computer Network and Multimedia Technology (CNMT '09)*, pp. 1–4, December 2009.
- [35] L. B. Yuste and H. Melvin, "Enhanced IPTV services through time synchronisation," in *Proceedings of the 14th IEEE International Symposium on Consumer Electronics (ISCE '10)*, pp. 1–6, June 2010.
- [36] I. Vaishnavi, P. Cesar, D. Bulterman, O. Friedrich, S. Gunkel, and D. Geerts, "From IPTV to synchronous shared experiences challenges in design: distributed media synchronization," *Signal Processing*, vol. 26, no. 7, pp. 370–377, 2011.
- [37] B. Cheng, "Virtual browser for enabling multi-device web applications," in *Proceedings of the Workshop on Multi-device App Middleware (Multi-Device '12)*, pp. 1–6, 2012.
- [38] N. Q. K. Duong, C. Howson, and Y. Legallais, "Fast second screen TV synchronization combining audio fingerprint technique and generalized cross correlation," in *International Conference on Consumer Electronics*, p. 4, 2012.
- [39] S. Soursos and N. Doulamis, "Connected TV and beyond," in *Proceedings of the Consumer Communications and Networking Conference*, pp. 582–586, 2012.
- [40] E. Gamma, R. Helm, R. Johnson, and J. Vlissides, *Design Patterns: Elements of Reusable Object-Oriented Software*, Addison Wesley, 1994.
- [41] A. López, M. Noè, and G. Fernández, "Linkage of additional contents to moving objects and video shots in a generic media framework for interactive television," in *The International Society for Optical Engineering*, vol. 5600 of *Proceedings of SPIE*, pp. 31–39, 2004.
- [42] T. A. Tavares, A. Burlamaqui, D. A. Da Silva Albino et al., "Sharing virtual acoustic spaces over interactive TV programs—presenting "virtual cheering" application," in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '04)*, vol. 3, pp. 2175–2178, June 2004.
- [43] M. Fink, M. Covell, and S. Baluja, "Social-and interactive-television applications based on real-time ambient-audio identification," in *Proceedings of the 11th edition of the European Interactive TV Conference series (EuroITV '06)*, p. 10, 2006.
- [44] R. Jaeger, P. Dunker, and A. Hanusch, "Enabling object based in-content placement for digital video television systems," in *Proceedings of the 8th IEEE International Symposium on Multimedia (ISM '06)*, vol. 2, pp. 662–666, December 2006.
- [45] E. Kawai, S. Yamaguchi, T. Akafuji, and K. Kandori, "A framework of a content distribution service synchronized with TV broadcasting programs," in *Proceedings of the 3rd IEEE Consumer Communications and Networking Conference (CCNC '06)*, vol. 2, pp. 1063–1067, January 2006.
- [46] J. Cha, Y. Seo, Y. Kim, and J. Ryu, "An authoring/editing framework for haptic broadcasting: passive haptic interactions using MPEG-4 BIFS," in *Proceedings of the 2nd Joint EuroHaptics Conference and Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems, World Haptics (WHC '07)*, pp. 274–279, March 2007.
- [47] F. P. Margalef, A. López, and G. Fernández, "Frame-accuracy synchronisation for mobile TV interactivity," in *Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB '09)*, pp. 1–5, May 2009.
- [48] T. W. L. Lai-Yeung and S. W. W. So, "TV as a multimedia synchronous communication for cooking and eating activities: analysis of TV cooking shows in Hong Kong," in *Proceedings of*

- the *IEEE International Symposium on Multimedia (ISM '10)*, pp. 302–307, December 2010.
- [49] L. B. Yuste, S. M. Saleh Al-Majeed, H. Melvin, and M. Fleury, “Effective synchronisation of hybrid broadcast and broadband TV,” in *Proceedings of the IEEE International Conference on Consumer Electronics (ICCE '12)*, pp. 160–161, January 2012.
- [50] G. Forbes, “Closed captioning transmission and display in digital television,” in *Proceedings of the International Workshop on Digital and Computational Video*, pp. 126–131, 2001.
- [51] J. Sung, D. Kim, Y. Doh, H. Lee, and J. S. Choi, “Broadcasting associated service architecture and its synchronization support,” in *Proceedings of the IEEE International Symposium on Consumer Electronics*, pp. 251–256, September 2004.
- [52] F. Hendrickx, T. Beckers, N. Oorts, and R. D. van Walle, “An integrated approach for device independent publication of complex multimedia documents,” in *Proceedings of the 9th IASTED International Conference on Internet and Multimedia Systems and Applications (IMSA '05)*, pp. 347–352, August 2006.
- [53] M. C. Marques Neto and C. A. S. Santos, “An event-based model for interactive live TV shows,” in *Proceedings of the 16th ACM International Conference on Multimedia (MM '08)*, pp. 845–848, October 2008.
- [54] J. Yun, H. Lee, and K. Park, “Orchestral media: the method for synchronizing single media with multiple devices for ubiquitous home media services,” in *Proceedings of the 3rd International Conference on Convergence and Hybrid Information Technology (ICCIT '08)*, vol. 1, pp. 340–345, November 2008.
- [55] M. C. Marques Neto and C. A. S. Santos, “An approach based on events for treating the late tuning problem in interactive live TV shows,” in *Proceedings of the ACM International workshop on Events in Multimedia*, pp. 41–48, 2009.
- [56] J. Yun, J. Jang, and K. Park, “Real-sense media synchronization technology based on the SMMD,” in *Proceedings of the International Conference on Consumer Electronics (ICCE '10)*, pp. 71–72, January 2010.
- [57] L. Lancercica, L. Dairaine, F. de Belleville, H. Thalmensy, and C. Fraboul, “MITv—a solution for an interactive TV based on IP multicast over satellite,” in *Proceedings of the IEEE International Conference on Multimedia and Expo (ICME '04)*, vol. 3, pp. 2159–2162, June 2004.
- [58] E. Papaioannou, K. Karpouzis, P. de Cuetos et al., “Melisa—a distributed multimedia system for multi-platform interactive sports content broadcasting,” in *Proceedings of the 30th EUROMICRO Conference*, vol. 30, pp. 222–229, September 2004.
- [59] P. Leroux, V. Verstraete, F. de Turck, and P. Demeester, “Efficient management of synchronised interactive services through the design of MCDP middleware,” in *Proceedings of the Australasian Telecommunication Networks and Applications Conference (ATNAC '07)*, pp. 215–220, December 2007.
- [60] M. F. Moreno, R. M. D. R. Costa, and L. F. G. Soares, “Sincronismo entre fluxos de mídia continua e aplicaçoes multimidia em redes por difusao,” in *Proceedings of the 14th Brazilian Symposium on Multimedia and the Web (WebMedia '08)*, pp. 202–209, 2008.
- [61] R. M. D. R. Costa, M. F. Moreno, and L. F. G. Soares, “Ginga-NCL: supporting multiple devices,” in *Proceedings of the 15th Brazilian Symposium on Multimedia and the Web (WebMedia '09)*, October 2009.
- [62] H. Lee, C. Yoon, and H.-W. Lee, “Smart Screen Service platform enabling target advertising,” in *Proceedings of the International Conference on Information and Communication Technology Convergence (ICTC '10)*, pp. 527–528, November 2010.
- [63] K. Matsumura, M. J. Evans, Y. Shishikui, and A. McParland, “Personalization of broadcast programs using synchronized internet content,” in *Proceedings of the International Conference on Consumer Electronics (ICCE '10)*, pp. 145–146, January 2010.
- [64] C. Howson, E. Gautier, P. Gilberton, A. Laurent, and Y. Legallais, “Second screen TV synchronization,” in *Proceedings of the 1st IEEE International Conference on Consumer Electronics (ICCE '11)*, pp. 361–365, Berlin, Germany, September 2011.
- [65] J. Lyle, S. Monteleone, S. Faily, D. Patti, and F. Ricciato, “Cross-platform access control for mobile web applications,” in *Proceedings of the IEEE International Symposium on Policies for Distributed Systems and Networks*, pp. 37–44, 2012.
- [66] F. M. Matsubara and M. Kawamori, “Lightweight interactive multimedia environment for TV,” *IEEE Transactions on Consumer Electronics*, vol. 57, no. 1, pp. 283–287, 2011.
- [67] A. Baba and K. Matsumura, “Seamless, synchronous, and supportive: welcome to hybridcast: an advanced hybrid broadcast and broadband system,” *Consumer Electronics Magazine*, vol. 1, no. 2, pp. 43–52, 2012.
- [68] C. Kohnen, C. Kobel, and N. Hellhund, “A DVB/IP streaming testbed for hybrid digital media content synchronization,” in *Proceedings of the International Conference on Consumer Electronics*, pp. 136–140, 2012.
- [69] L. B. Yuste and H. Melvin, “Client-side multisource media streams multiplexing for HbbTV,” in *Proceedings of the International Conference on Consumer Electronics*, pp. 1–5, 2012.
- [70] A. L. S. Gradvohl and Y. Iano, “An approach for interactive television based on insertion of hypermedia information in MPEG standard video,” in *Proceedings of the IEEE International Symposium on Consumer Electronics*, pp. 25–30, September 2004.
- [71] H. M. Stokking, M. O. van Deventer, O. a. Niamut, F. a. Walraven, and R. N. Mekuria, “IPTV inter-destination synchronization: a network-based approach,” in *Proceedings of the International Conference on Intelligence in Next Generation Networks*, pp. 1–6, 2010.
- [72] F. Boronat, M. Montagud, H. Stokking, and O. Niamut, “The need for inter-destination synchronization for emerging social interactive multimedia applications,” *Communications Magazine*, vol. 50, no. 11, pp. 150–158, 2012.
- [73] M. Montagud and F. Boronat, “Enhanced adaptive RTCP-based Inter-destination multimedia synchronization approach for distributed applications,” *Computer Networks*, vol. 56, no. 12, pp. 2912–2933, 2012.
- [74] M. Montagud, F. Boronat, H. Stokking, and R. Brandenburg, “Inter-destination multimedia synchronization: schemes, use cases and standardization,” *Multimedia Systems*, vol. 18, no. 6, pp. 459–482, 2012.
- [75] B. Li, J. H. Errico, H. Pan, and I. Sezan, “Bridging the semantic gap in sports video retrieval and summarization,” *Journal of Visual Communication and Image Representation*, vol. 15, no. 3, pp. 393–424, 2004.
- [76] R. G. Cattelan, C. Teixeira, R. Goularte, and M. D. G. C. Pimentel, “Watch-and-comment as a paradigm toward ubiquitous interactive video editing,” *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 4, no. 4, article 28, 2008.
- [77] A. Crivellaro and M. P. de Almeida, “The real time audience system, one minute or nothing,” in *Proceedings of the 4th*

- International Conference on Automated Solutions for Cross Media Content and Multi-Channel Distribution*, pp. 204–207, November 2008.
- [78] J. E. Garcia, A. Ortega, E. Lleida, T. Lozano, E. Bernues, and D. Sanchez, “Audio and text synchronization for TV news subtitling based on automatic speech recognition,” in *Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB '09)*, May 2009.
- [79] V. G. Motti, R. Fagá, R. G. Catellan, M. Da Graça C. Pimentel, and C. A. C. Teixeira, “Collaborative synchronous video annotation via the watch-and-comment paradigm,” in *Proceedings of the 7th European Conference on European Interactive Television Conference (EuroITV '09)*, pp. 67–76, June 2009.
- [80] J. Gao, Q. Zhao, and Y. Yan, “Automatic synchronization of live speech and its transcripts based on a frame-synchronous likelihood ratio test,” in *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP '10)*, pp. 1622–1625, March 2010.
- [81] M. de Castro, D. Carrero, L. Puente, and B. Ruiz, “Real-time subtitle synchronization in live television programs,” in *Proceedings of the IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB '11)*, June 2011.
- [82] D. Das, P. Sinha, A. Ghose, and C. Bhaumik, “An interactive system using digital broadcasting and quick response code,” in *Proceedings of the 15th IEEE International Symposium on Consumer Electronics (ISCE '11)*, pp. 397–400, June 2011.
- [83] F. L. S. Ferreira, F. H. Lemos, G. P. B. Neto, T. M. U. de Araujo, and G. L. D. S. Filho, “Providing support for sign languages in middlewares compliant with ITU J.202,” in *Proceedings of the 13th IEEE International Symposium on Multimedia (ISM '11)*, pp. 149–156, December 2011.
- [84] I. Bisio, A. Delfino, G. Luzzati et al., “Opportunistic estimation of television audience through smartphones,” in *Proceedings of the International Symposium on Performance Evaluation of Computer and Telecommunication Systems*, pp. 1–5, 2012.
- [85] S.-H. Dai, Z.-Y. Zhu, and G.-Y. Lin, “Synchronization model design and implementation of distributed multimedia system,” in *Proceedings of the 3rd International Conference on Innovative Computing Information and Control (ICICIC '08)*, p. 209, June 2008.
- [86] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, “RTP: a transport protocol for real-time applications,” RFC 3550, 2003.
- [87] G. Harboe, N. Massey, C. Metcalf, D. Wheatley, and G. Romano, “The uses of social television,” *Computers in Entertainment*, vol. 6, no. 1, article 8, 2008.



**Hindawi**

Submit your manuscripts at  
<http://www.hindawi.com>

