# Social Media Usage Combined with TV/Video Watching: Opportunities and Associated Challenges

## Towards Augmented, Interactive, Personalized and Shared Experiences

Mario Montagud[1,2], Pablo Cesar
[1] Distributed and Interactive Systems (DIS) Group
Centrum Wiskunde & Informatica (CWI)
Amsterdam, The Netherlands
montagud@cwi.nl, P.S.Cesar@cwi.nl

Fernando Boronat, Dani Marfil
[2] IIM R&D Group, Communications Department
Universitat Politècnica de València (UPV)
Grau de Gandia (València), Spain
fboronat@dcom.upv.es, damarre@upv.es

***Abstract*—This paper provides an overview of the impact and opportunities provided by Social Media and other social interaction tools when watching TV/video content. The analysis has been conducted from the viewpoints of both individual and shared media experiences between remote users. On the one hand, many possibilities provided by Social Media when each user is concurrently consuming multiple related media content, either on a single device or on different devices (e.g., in multi-screen scenarios), are described. On the other hand, the potential of Social Media and other interaction tools when multiple remote users are concurrently consuming the same media content (e.g., in Social TV) is discussed. In addition, the paper highlights some remaining challenges and open issues that need to be addressed in the near future to truly provide augmented, interactive, personalized and shared experiences, combining Social Media usage and TV/video content consumption. Finally, as a real use case, the social interaction, presence and privacy mechanisms provided by a media sharing platform under development, called Wersync, are described.**

***Keywords—Social Media, TV Watching, Social TV, Multi-Screen Scenarios, Interaction, Media Sync, Shared Experiences***

## I. INTRODUCTION

Nowadays, we are witnessing a dual paradigm shift regarding media (and specifically TV) content consumption. On the one hand, from the viewpoint of individual media experiences, the traditional setting in which each user was passively consuming a single media content on a unique device (e.g., lean-back TV viewing) is being replaced by a more active setting, in which a rich variety of media content (e.g., audio, video, textual information…), delivered via either the same or different (broadband and broadcast) technologies and sources, can be simultaneously consumed either on the same device or on different, but close-by, devices (e.g., connected TVs, tablets, smartphones…) in multi-screen scenarios. On the other hand, from the viewpoint of group shared (or collective) media experiences, the traditional setting in which various users gather at a specific location for consuming media together (e.g., watching TV) can be virtually recreated when the users are far apart, thanks to the advances in ubiquitous media delivery technologies together with the proliferation and improvement of connected consumption devices. This allows geographically distributed users to concurrently consume the same media content while socially interacting, as if they were co-located.

This article provides an overview, from these two viewpoints, of the impact and opportunities provided by Social Media and social interaction tools when watching TV/video content, such as: i) accessing to extra related content; ii) adding extra functionalities/services (e.g., votings, TV quizzes…); iii) fostering social interaction; iv) providing a feeling of *"networked togetherness"* (i.e., a feeling of co-presence and physical togetherness when the involved users are far apart and communicating via networked services); and v) empowering the audience's engagement. In addition, it highlights some existing challenges and open issues that require further research to provide truly enriched, personalized, interactive and shared media experiences, combining Social Media and social interaction tools with TV/video content consumption. Finally, as real use case, the social interaction, presence and privacy features that are currently supported by a media sharing platform under development, called Wersync, are described.

## II. SOCIAL MEDIA IN INDIVIDUAL MEDIA EXPERIENCES

Given the multiple media delivery and consumption possibilities at users' disposal, significant efforts have been devoted towards achieving a coordination and convergence between the different technologies for consuming media. The goal is to conceive them as a whole, rather than as isolated worlds. A piece of evidence is the recent Hybrid Broadcast Broadband TV (HbbTV) 2.0 standard [1], which provides the mechanisms for harmonizing the delivery and consumption of interactive broadcast and broadband TV-related content through connected TVs and secondary (or companion) devices.

This emerging and converged hybrid broadcast and broadband media landscape, in combination with the popularization and mass adoption of Social Media, opens the door to a new wide range of extra functionalities and business models regarding media consumption. Within the TV scope, it allows offering innovative forms of augmented, interactive and personalized services to the audience, which significantly contributes to empower their immersion and engagement. A direct consequence of this multi-connected media ecosystem is

a transformation of the users' TV viewing habits. Users no longer just (passively) watch a single TV event. Instead, they increasingly tend to (actively) multi-task, by using their secondary devices to concurrently consume additional media content when watching TV.

Various surveys (e.g., see [2, 3]) have investigated this paradigm shift from single-screen, passive and unidirectional media experiences towards multi-screen, proactive and interactive media experiences, in which broadcast, broadband and Social Media converge. It has been found that a big percentage (up to around 80% in specific cases [2]) of consumers use a secondary device while watching TV. Although the multi-tasking activity can be completely unrelated to the TV watching experience, such as when users are surfing the web or e-mailing, a wide variety of new related media possibilities and services can be exploited. In particular, the use of Social Media (through either the main screen or secondary screens) when watching TV can provide an augmented experience in which users can discover extra information about the media content being consumed, and can be provided with a rich set of additional (interactive) possibilities.

Through Social Media channels, the audience can be provided by additional information about the TV content, which, in turn, can be provided by either the content provider or by other consumers. Apart from textual information (e.g., chat messages, news, status updates, statistics…), Social Network platforms can also convey pictures, short media clips (e.g., replays of video scenes or video comments via Vine app when using Twitter), advertisements and even URLs to additional related content (e.g., to related websites or media content). When using Twitter, such information can be retrieved by accessing to a targeted profile/account or by listening to a specific hashtag, while when using Facebook it can be done, for instance, by accessing to a specific page (e.g., the official page of a content provider or of a sports club). As examples, a user watching a film can get additional information about the actors, other films by the same director, or even access to additional video scenes or histories not included in the film. Additionally, users can access to the comments by people of interest (e.g., their favorite sportscaster, the protagonists of the TV content…) or by just other consumers, be aware of the existence of trends (e.g., trending topics) about the TV event and even they have the opportunity of actively participating in the TV event by posting comments in these Social Media channels. What is more, users also have the chance of voting, rating aspects, answering to quizzes (so that they can compete and win prices) and even participating in bets (e.g., in sports events). For example, when using Twitter, users can easily vote (e.g., for the man of the match in a sports event) between two options via the *Retweet* and *Favorite/Love* controls or between many of them by using specific hashtags, and between three options via the *Comment*, *Like* and *Share* controls when using Facebook.

All these possibilities provided by Social Media contribute to provide augmented, interactive and personalized TV services, and are very relevant to: i) allow new media business and advertising models; ii) increase the audience's immersion and engagement; and iii) encourage the participation of new consumers in TV events. As a proof of evidence, one third of the respondents in one of the surveys summarized in [2] stated they were more likely to watch TV shows if there is a huge social buzz around it.

Likewise, it is important to mention two issues. First, most of the previous features, and additional interactive ones, are also commonly provided by customized TV companion apps in the secondary devices, which, in turn, can also integrate Social Media functionalities. Second, not all the secondary device users want to be automatically provided by extra information, but they prefer to search it by themselves (e.g., by visiting related websites), as it has been reflected in the surveys conducted within the umbrella of a Spanish research project about hybrid media consumption the authors are immersed in.

## III. Social Media in Shared Media Experiences

Certain traditional forms of media consumption involve social interaction between users. For instance, family members or friends often gather at a single location for consuming media (e.g., for watching TV) together. The typical scenario is a group of friends watching a live football match at a friend's home. Actually, the shared consumption of media is frequently the catalyst why the users meet up, as it allows discussing about common interests, re-living experiences and sharing emotions, thus leading to rich human interactions, increasing the users' engagement and even strengthening the social bonds ([4], [5]).

Unfortunately, many times, a myriad of practical factors prevent from a physical congregation. In the current global society, relatives and friends live in, or frequently move to, different geographical locations (cities, countries or continents) due to various reasons (e.g., studies, job, business, vacations…). This geographical segregation has arisen the need for recreating such shared media experiences between remote users, as they are still interested in remaining connected and in socializing with their friends and relatives living apart. As mentioned earlier, this transition from physical togetherness towards "*networked togetherness"* around media content is becoming a reality thanks to the latest advances in media delivery technologies and the proliferation of connected devices, but, additionally (and as in the previous section), a third factor plays a key role: the mass adoption of Social Media.

At present, the use of Social Media for expressing our opinions and thoughts about almost everything, and for communicating with other online users, has become a routine activity. If we bring it to the TV consumption area, apart from the functionalities that have been cited in the previous section, Social Media also contributes to the creation of live discussion and interaction forums when the same media content is concurrently consumed by remote users. Indeed, recent studies ([2]) have shown that around 40% of viewers using Social Media (through secondary devices) during TV consumption are frequently discussing about what they are actually watching. These communications and interactions allow emulating a feeling of *"networked togetherness"*.

The combination of social networking functionalities with simultaneous consumption of TV content is commonly referred

to as Social TV (a.k.a. *"watching apart together"*), which aims at connecting and fostering real-time interaction between TV viewers, thus transforming the TV watching experience into a social event [5, 6]. For instance, the co-located friends in the above example can now watch the football match each from their own home, while being able to converse, discuss about its evolution, and cheer together (e.g., when goals are scored). An overview and categorization of the Social TV research area is provided in [6].

The communications and social interactions between the involved TV viewers can occur either through the built-in IP communication features of the connected TVs or through the personal secondary devices. Likewise, different communication modalities can be used, such as text chat tools and audio/video conferencing services (and combinations thereof). In the next section, some implications regarding the use of the different social interaction modalities and the creation of online communities are discussed.

## IV. Challenges and Open Issues

This Section provides a discussion about some existing challenges and open issues that need to be overcome to enable truly enriched, personalized and shared experiences, combining TV/video consumption and Social Media usage. Other relevant challenges ([7, 8]), such as Quality of Service (QoS) provisioning, Quality of Experience (QoE) assessment, design of efficient media adaptation and delivery methods, heterogeneity and inter-operability issues, and cross-domain session handling, have not been considered, because of a less direct relationship with the analysed topic.

### A. Interactivity & Synchronization

Two key related requirements must be met in these kinds of augmented and shared media services. First, the end-to-end delivery delays for all the involved media streams and destinations must be kept within tolerable bounds in order to enable truly interactive services and natural communications. Second, different types of media synchronization (sync hereafter) must be provided:

i) *intra-media sync*: sync of the media playout for each involved media element (e.g., audio, video, textual information…) in order to achieve a natural and smooth playout, despite of the existence of jitter;

ii) *inter-media sync*: sync of the media playout between different media elements (e.g., lip-sync) within specific devices, in order to preserve their temporal dependences and equalize delay differences between them. These media elements can be sent in different streams or in a multiplexed stream.

iii) *Inter-Device Sync (IDES) / Inter-Destination Media Sync* (IDMS): sync of the media playout of the same or different media elements across devices. If the involved devices are close-by (e.g., multi-screen scenarios), the term IDES is commonly used, and if the involved devices are geographically distributed (e.g., in different houses, cities, countries…), the term IDMS is commonly used.

When the different media elements to be synchronized, either on the same devices or across devices, are sent via different delivery protocols or technologies, such as broadcast and broadband (e.g., in HbbTV), the term *hybrid sync* is commonly used

In shared media experiences (e.g., in Social TV), it is especially relevant the compensation of the delay differences between the involved media elements (i.e., achieving inter-media sync), devices (i.e., achieving IDES) and destinations (i.e., achieving IDMS), especially when the involved users are discussing and socially interacting within the context of the content being consumed, to guarantee that all of them perceive the same events at (approximately) the same time. Otherwise, the absence of the above kinds of media sync can lead to incoherent interactions and to frustrating situations, such as users being aware of a goal through the cheering of a friend via the chat channel before actually watching/hearing it through the local consumption device(s). In these scenarios, it is also necessary the minimization of delays for all the involved content streams and devices.

Up to date, many proprietary and standard solutions have been devised for each one of the above media sync types (see e.g., [9, 10]), especially for audio and video communications. However, it is very challenging to provide sync between the multiple (time-sensitive) media information (probably of different nature: audio, video, text-based…) provided by different, and independent, sources. To achieve it, it is necessary to include capturing/transmission timestamps [11], obtained from either synchronized or traceable clocks, within the involved content streams so they can be time aligned at the receiver side and an overall consistent view can be presented to the users.

The insertion and interpretation of timestamps in ad-hoc content ingest platforms is not a big challenge, but doing so when using third-party platforms and off-the-shelf equipment becomes much more complicated. This is specially an issue for the messages from *"public"* Social Media platforms. Moreover, these timestamps will need to remain unchanged through the end-to-end delivery chain, or some mapping mechanisms should be provided if they are replaced (e.g., because of transcoding).

Besides, the end-to-end delay of messages from Social Networks plays a key role in these kinds of experiences. On the one hand, if a message is presented to the users earlier than the associated media content, it can spoil the experience. On the other hand, if the messages arrive too late, then they may be no longer relevant. Another (more uncontrollable) issue is that the timestamps of these messages (if included) will relate to its transmission instant, but not the event being perceived and described in these messages, and different users may need different time to write a specific message (it will depend on their typing ability and on how long the message is).

In this context, it is worth to mention the research recently conducted within the umbrella of the FP7 STEER project (http://fp7-steer.eu/) [12], which aimed at augmenting live broadcasted events (e.g., a sports event, a concert…) with live user-generated content in a synchronized manner. This user-generated content consists of both video content, recorded by

users with their personal devices (e.g., smartphones), and Social Media content (e.g., Facebook posts and Twitter feeds), provided by users visiting the event (e.g., users attending a sports event at the stadium) and by other users remotely watching the event (e.g., at home). These kinds of scenarios, combining networked media, social networks and user-generated content are defined as *"Social Telemedia"* services. Herein, the sync between all the involved streams across all the involved destinations imposes even more stringent challenges, because the information from Social Media platforms comes both from users who are actually physically attending the event and from users who are remotely watching the event via networked technologies. In such a case, it is necessary to discern between both types of users, because the local users perceive the event in real-time, while the remote users perceive the event a few seconds (or even minutes) later, due to the end-to-end media delivery delay. A potential solution is to differentiate between real-world timestamps and "consumed media content" timestamps, so that the timing of the messages will be in correlation with the timing of the event. In any case, it is very challenging, and almost impossible, to provide consistent interactions in case of high end-to-end delivery media delays.

Moreover, regarding the use of (real-time) social interaction channels in group shared experiences, the work in [13] identifies and discusses various limitations and constraints of using Facebook and Twitter, such as low flexibility for embedding and retrieving synchronization metadata (e.g., timelines), high dependence on third-party components and infrastructure, non-guaranteed scalability and availability (e.g., bounds in the traffic volume and/or rate per period of time), high end-to-end delays (i.e., delays between the instant at which a message is posted and the instant it is presented to the users), and need for filtering and refresh mechanisms. However, the usage of *"public"* Social Networks as the social interaction channels in shared media experiences is not the unique solution, but other alternative ad-hoc chat tools with lower latency and higher flexibility (e.g., in terms of maintenance, for creating independent chat rooms for different groups of users, for including synchronization metadata…), can be developed. For instance, one alternative is to use the standard Extensible Messaging and Presence Protocol (XMPP), as in [4], and other alternative is to use Javascript components, such as *Node.js* and *Socket.IO*, as in [14].

### B. Scalability & Privacy

The amount of data related to TV events from social networks can be immensely large. When using Facebook, users can retrieve and post messages from/to official pages (e.g., managed by the content providers or fans club), or can explicitly create new groups to enable more controlled and semi-private discussion forums and chat rooms for smaller and ad-hoc groups.

When using Twitter, users can also retrieve information from official profiles/accounts, but they can also retrieve and post messages via hashtag-based filtering mechanisms. In addition, other more sophisticated mechanisms can be provided. For instance, it is possible to provide aggregated or filtered information, such as statistics, frequently used terms or tweets from a curated set of influential Twitter users. It is also possible to provide social analytics functionalities, such as event profiling and tracking, sentiment mining, recommendations, and user authority measurement. Regarding this last feature, it is possible to categorize or rank the messages based on the "authority / relevance" levels of their users (e.g. by taking into account their number of friends or followers, or the number of likes, replies, favorites and retweets to their messages). Likewise, the ranking mechanisms can be personalized (e.g., by considering more relevant the messages from the users' contacts).

However, all these filtering and aggregation mechanisms will add extra delays due to the associated processing and the overall view construction processes.

All the scalability and privacy (and interactivity) issues that the use of Social Networks may involve (with the exception of creating Facebook groups), can also be overcome with the use of ad-hoc chat channels (as previously mentioned), but with the drawback of having to add and employ other tools than the ones customers are used to.

### C. Dynamic Building of Social Communities

TV events can involve a large amount of viewers and Social Media users. In these situations, it is quite common that specific groups of known users (e.g., family members, friends…) create their own interaction groups (e.g., Facebook groups, WhatsApp groups, or Skype calls). However, it is also possible to dynamically create and manage ad-hoc social communities, based either on the explicit feedback provided by users or on implicitly collecting and analyzing their activity (e.g., content being consumed), interactions, comments, attitudes, preferences (likes, favorite tweets, retweets…), feelings, interests (e.g., number of friends, accounts being followed…) and profiles (e.g., demographics, age, gender, occupation…) in Social Media platforms. The creation of communities can be assisted by recommendation systems and the use of advanced algorithms (e.g., sentiment mining, emotion detection…), and can be seen as *"micro social networks"* grouping users with common aspects, thus providing them the content and features they are most interested in (e.g., content syndication, personalized advertisements) and fostering more tailored socialization around media content.

The relevance of *"community building"* requirements to allow the structuration of large amounts of users when commenting and discussing about TV content was highlighted in [6], and has been also considered in recent EU projects, such as in iNEM4YOU (Interactive Networked Experiences in Multimedia for You) [7] and in the ongoing SAM (Socializing Around Media) project (http://samproject.net).

### D. Social Interaction Modalities

As mentioned, three social interaction modalities can be used in shared media experiences: text, voice and video chat (and combinations thereof). Previous studies have explored the influence of the use of these interaction channels on the tolerable asynchrony thresholds (i.e., delay differences) in Social TV scenarios [15]. However, as far as we know, no

surveys and experimental studies have been conducted to analyze the users' preferences, and the impact on the users' perceived QoE (e.g., usability of the system, feeling of togetherness, level of engagement, naturalness of interactions, distraction, comfortability…) when using each variant of the communication channel or different combinations of them. This is one of the research goals of the Spanish project we are currently immersed in.

### E. Natural and Immersive Experiences

Although it has not been demonstrated through experimental studies yet, one can guess that video chat channels provide more natural, interactive, immersive and comfortable communications than text chat channels. However, the integration of video conferencing services in shared media experiences still faces many challenges and open issues. Apart from media sync, scalability, cross-domain session handling and dynamic creation of ad-hoc social communities, the work in [16] (done within the scope of EU VCONECT project, http://www.vconect-project.eu) calls on the research community to focus on other fundamental aspects. The idea is to move beyond the "talking heads" paradigm that characterizes most (multi-party) video communication systems, in which users have a static and quite inflexible face-to-face communication, to more advanced systems that must not only be able to dynamically adapt to changes at the network layer (e.g., congestion, increase of delays…) and at the application layer (e.g., dynamic sessions, turn-taking…), but must also be able to understand the nature of the shared experience and the social layer of interactions. To realize this vision, it will be necessary to take intelligent decisions and dynamic/seamless adaptations (e.g., encoding and delivery methods, screen layout…), based on several technological and perceptual aspects, such as contextual information (e.g., the communication and interaction context, the participants' roles, the conversation dynamics, the strength of inter-personal ties…), cue processing (e.g., facial expressions and body language could be captured to infer conversation patterns and interaction quality among participants), number of participants, QoS/QoE monitoring, etc.

Solving these challenges will open new opportunities for the realization of high-quality (context-aware and) socially-aware video communications, enabling truly connected, immersive and natural shared experiences that can more efficiently convey emotions and provide stronger feelings of networked togetherness.

### F. Bandwidth Limitations

These augmented and shared media experiences may currently suffer from bandwidth limitations, depending on the amount of media content being consumed, and the characteristics and conditions of the access networks being used. However, it is expected that the developments in encoding and delivery techniques, the ever increasing bandwidth capacity of core and access networks, in addition with the deployment of next-generation technologies, such as 5G, will contribute to overcome these limitations, enabling a more efficient support for these experiences and for device-to-device communications.

## V. The Use Case of Wersync

As a related use case, this Section describes the Social Media and social interaction features provided by Wersync [14], which is an adaptive web-based platform for distributed media consumption and social interaction across remote users, being developed at Polytechnic University of Valencia (UPV). Wersync allows the creation of independent groups of users, each of which being able to consume the same or different media content in a synchronized manner. When accessing to the platform, each user can choose between creating a new shared session (by also selecting the clip to be watched from an online video library) and joining an on-going one. Wersync provides two main interaction mechanisms. First, it allows sharing the navigation control (i.e., *play*, *pause* and *seek to*) commands of the media player between all the users in a shared session. Second, it provides two modalities of text chat channels for users' interaction. The first one is based on integrating Twitter via its Javascript API. The drawback is, as previously mentioned, the interactivity limitations and the non-instantaneous refresh of the timeline. Moreover, the use of Twitter involves having a "public" chat room, even though some filtering mechanisms can be used, by listening on a specific hashtag (e.g., *#Wersync*, or even adding the session id as a suffix). The second one is an ad-hoc text chat tool, developed by using the Javascript *Socket.IO* library. It provides much better performance in terms of delays, and much more flexibility for inserting and interpreting timestamps within the chat messages, thus allowing their time-alignment with the other media components in the shared session (i.e., inter-media sync), than the use of Twitter. Moreover, unlike Twitter, it also allows having private chat rooms for each shared session. The surveys and interviews we have conducted with users so far do not reveal significant differences between their preferences regarding the use of each one of the above modalities. That is the reason we decided on integrating both of them in Wersync. Moreover, Twitter is used for another purpose, as explained next.

A third form of interaction is currently under development, which consists of adding video conferencing services as the chat channel between the participants in each shared session. It is expected that this will provide a more comfortable, natural and realistic (face-to-face) interaction between users.

Wersync also provides two *"social presence"* mechanisms. The first one is an internal menu with drop-down lists, indicating the list of active sessions, their members and the media being consumed. This way, newcomers can check if they want to join any of the on-going sessions. The second one is an external presence mechanism, which consists of automatically posting a tweet every time a user creates or joins a session on Wersync (if he/she is logged in on Twitter and gives his/her consent for that). This tweet will include the appropriate information to univocally identify the shared session, including the user's nick in Wersync, the clip being consumed, hashtags (e.g., *#Wersync*, *#user_nick* and *#session_id*) and a URL to join the shared session (see Figure 1). This announcement will allow external users to know about the activity of their Twitter contacts in Wersync, which will undoubtedly contribute to encourage their participation in on-going shared sessions. Additionally, the availability of

audiovisual chat channels will (implicitly) provide a third social presence mechanism.


Fig. 1. Tweet informing about the activity of a user in Wersync

Finally, Wersync provides two privacy mechanisms. First, despite of the Twitter notifications, the participation of new users in on-going shared sessions can be controlled. When each user requests to join a session, a message will be sent to the master/manager of that session, who can accept or reject that request. Second, the chat messages can be encrypted (if desired).

Prior to the development of Wersync, a survey was conducted to gain insight about the users' habits regarding sharing media experiences with remote users, their interest in the availability of platforms to efficiently support this, and their preferences and expectations regarding the functionalities to be provided by these platforms. The obtained results in that study clearly revealed the usefulness of, and the users' interest in, this platform, and its findings helped us in developing the functionalities that were demanded (e.g., the use of an ad-hoc chat tool in combination with Twitter, the addition of video chat tools…). The results of this study will be published together with the QoE evaluation of the platform.

## VI. Conclusions

This paper has provided an overview about the impact and potential opportunities offered by Social Media and social interaction tools when consuming TV/video content, from two different point of views: i) when isolated users are consuming multiple related content on either the same or on different (close-by) devices (e.g., in multi-screen scenarios); and ii) when multiple geographically distributed users are concurrently consuming the same media content, while socially interacting. In both kinds of scenarios, Social Media can contribute to: i) access to extra related content; ii) add extra functionalities/services (e.g., votings, TV quizzes…); iii) foster social interaction; iv) provide a feeling of *"networked togetherness"*; and v) empower the audience's immersion and engagement. All these features support the important role of Social Media in the current dual transition from passive, single-content and single-device TV/video experiences towards active, multi-content and multi-device TV/video experiences, and from physically shared TV/video experiences towards remotely shared TV/video experiences. In addition, this article has highlighted some remaining challenges and open issues that still need to be addressed in the near future to provide truly augmented, personalized, interactive, immersive and shared experiences, combining Social Media and social interaction tools with TV/video consumption. From these challenges, media sync (in all its facets) becomes a key requirement rather than a simple desire.

## Acknowledgment

The work by Mario Montagud has been carried out during the tenure of an ERCIM 'Alain Bensoussan' Fellowship Programme. UPV work has been funded, partially, by the "Fondo Europeo de Desarrollo Regional (FEDER)" and the Spanish Ministry of Economy and Competitiveness, under its R&D&I Support Program, in project with reference TEC2013-45492-R.

## References

[1] HbbTV Hybrid Broadcast Broadband TV version 2, ETSI TS 102 796 v1.3.1, October 2015, available at hbbtv.org.

[2] Social TV and second-screen viewing: the stats in 2012 www.theguardian.com/technology/appsblog/2012/oct/29/social-tv-second-screen-research, October 2012 (last access in October 2015).

[3] Nielsen, Social Media + Linear TV = Engaged Viewers, http://www.nielsen.com/us/en/insights/news/2015/live-tv-social-media-engaged-viewers.html, June 2015 (last access in October 2015).

[4] M. Wijnants, J. Dierckx, P. Quax, W. Lamotte, "synchronous MediaSharing: social and communal media consumption for geographically dispersed users", ACM MMSys 2012, February 2012.

[5] C. Timmerer, B. Rainer, "The Social Multimedia Experience", IEEE Computer, 47(3), pp. 67-69, March 2014.

[6] P. Cesar, D. Geerts, "Past, present, and future of social TV: A categorization", IEEE Consumer Communications and Networking Conference (CCNC), Las Vegas (USA), January 2011.

[7] C. Hesselman, et al., "Sharing enriched multimedia experiences across heterogeneous network infrastructures", IEEE Communications Magazine, 48(6), pp.54-65, June 2010.

[8] I. Vaishnavi, et al., "From IPTV to synchronous shared experiences challenges in design: Distributed media synchronization", Signal Processing: Image Communication, 26(7), pp. 370-377, August 2011.

[9] F. Boronat, J. Lloret, and M. García, "Multimedia group and inter-stream synchronization techniques: A comparative study", Information Systems, 34(1), pp. 108-131, March 2009

[10] M.O. van Deventer, H. Stokking, M. Hammond, J. Le Feuvre, and P. Cesar, "Standards for Multi-stream and Multi-device Media Synchronization", IEEE Communications Magazine, Accepted for Publication (in August 2015).

[11] L. Beloqui, F. Boronat, M. Montagud, H. Melvin, "Understanding Timelines within MPEG Standards", IEEE Communications Surveys and Tutorials, Accepted for Publication (in September 2015).

[12] H. Stokking, R. Kaptein, A. Veenhuizen, M. Spitters, O. Niamut, "Real Time Synchronization of Live Broadcast Streams with User Generated Content and Social Network Streams", MediaSync Workshop 2013, Nantes (France), October 2013

[13] M. Wijnants, P. Quax, W. Lamotte, "Cost-effective Web-based Media Synchronization Schemes for Real-time Distributed Groupware", WEBIST 2013, Auchen (Germany), May 2013.

[14] J. Belda, M. Montagud, F. Boronat, M. Martínez, J. Pastor, "Wersync: A Web-based Platform for Distributed Media Synchronization and Social Interaction", ACM TVX 2015, Brussels (Belgium), June 2015.

[15] D. Geerts, et al., "Are we in sync?: synchronization requirements for watching online video together", ACM CHI 2011, Vancouver (Canada), May 2011.

[16] P. Cesar, R. Kaiser, and M.F. Ursu, "Toward Connected Shared Experiences", IEEE Computer, 47(7), pp. 86-89, 2014.