An Architecture of Video Object Annotation Platform for Interactive TV

Hyun Namgoong, Kang Yong Lee, Kee Seong Cho, Won Ryu
Smart Screen Convergence Research Department
Electronics and Telecommunications Research Institute
Daejeon, South Korea
{nghyun, kanglee, chokis, wlyu}@etri.re.kr

Hong-Gee Kim
Biomedical Knowledge Engineering Laboratory
Seoul National University
Seoul, South Korea
hgkim@snu.ac.kr

***Abstract*— Current advances of television systems provide interactive user experiences. This paper proposes a Video Object Annotation Platform (VOAP), which enables TV viewers can be provided with additional information during they is watching TV. The platform anchored to semantic data serialization to support a sharing of annotations located to specified object at a moment. With the platform, content providers can publish additional information about objects appeared in a program and the viewer can easily access the information. We expect that the proposed platform will provides a novel business model for advertisement market by providing seamless information transfer among consumer (viewer), advertisers (company) and mediation (broadcasting company). As an ongoing work proposing an information exchanging environment in interactive TV platforms, we implemented a user-side application proving the proposed concept.**

Keywords-IPTV; Interactive TV; Video Object Annotation; Annotation Platform.

I. INTRODUCTION

With the name of IPTV (Internet Protocol Tele-Vision), current advances of TV systems provide people interactive experiences like VOD (Video On Demand) service and time-shifted watching [1]. Broadcasting companies and research groups are trying to put more services showing advantage of such interactive TV system [2], so that, the television viewers can enjoy various experiences in the TV screen. A TV answering user questions is not a new research topic, because, it ensures many new opportunities in commercial Ads. From casual one like "Who is that actor?" to complicated ones " Where the movie is taken from?" and " Where can I buy that thing?" the informative questions are easily come up when we watch TV. An environment to provide such additional information to viewers during watching TV is also important to broadcasting companies or content providers. They can not only answer such question with additional information, but also expose them to commercial advertisement at the screen without troubling them.

This paper proposes a video object annotation platform for digitalized and bi-directed TV environment, and it is anchored to current data serialization techniques that let data to be interchangeable and reusable through a standardized data representation format like RDF (Resource Description Framework) [RDF refer]. The proposed platform support a sharing of annotations located to specified object at a moment, so that, content providers publishes additional information related to program. Then, the platform ensures that viewer of the movie or TV series can see and retrieve the information in the interactive TV.

We expect that the proposed platform will provide a novel business model for Ad market by providing seamless information transfer among consumer (viewer), advertisers (company) and mediation (broadcasting company). Contributions of this paper can be summarized as like; 1) a data schema for sharable movie annotation represented, and 2) a technical design of movie annotation platform proposed.

The rest of this paper is organized as follows. Our basic idea on movie annotation and the other related work are described in Section 2. Overall architecture of Video Object Annotation Platform (VOAP) will be described in Section 3 with two subsections, which explain details of key elements of the platform. Finally, Section 4 concludes the paper.

II. MOVIE ANNOTATION

Video annotation is tasks of associating graphical objects with moving objects on the screen. In existing interactive applications, only still images can be annotated, as in the "telestrator" system [3] used in American football broadcasting. Jeroen [4] developed an adaptive movie annotation with speech recognition techniques. The system helps people easily attach annotation by mapping scripts and a specific moment which the scripted is spoken. In the system the additional information like "who?" and "where" can be obtained from script of the program. A system proposed in [5, 6] helps easy annotating on moving objects in the scene by human. The system supports descriptive labels, illustrative sketches, thought and word bubbles communicating speech or intent, path arrows indicating motion, and hyperlinked regions.

The previous works mostly are focused on how we can easily attach annotations on the movies and how a movie is explained by annotations on divided scenes. Our work is focused on the annotations attached to appeared object at each moment in a movie and how the annotations can be shared and spent by viewers. Of course, a technology for easy and correct creation of annotations is an essential issue

 ISBN: 978-1-61208-215-8

to be studied. We rather assume that the issues of creation of annotations could be solved through human labors if we have useful business model or motivation as we can see the Wikipedia [7] case. Therefore, our work is more focused on how the annotations can be used to enhance information viewers and content providers or to provide useful application.

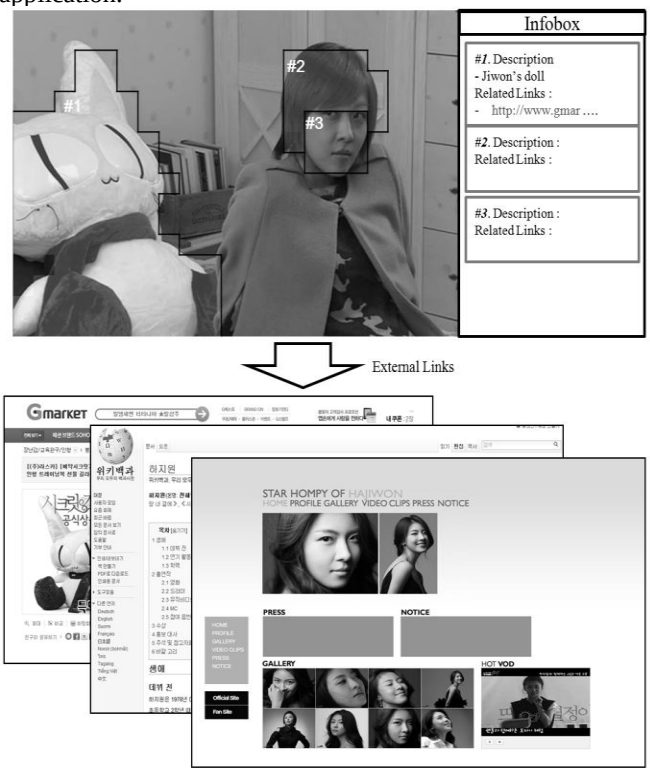


Figure 1. Examplied VOAP application running on IPTV

Figure 1 shows an example use of the attached annotation on the TV program. The annotation includes simple description about the objects we can currently see on the screen. Also, they provide external links containing useful information about the objects. By watching TV, a viewer can make a pause for a moment to have a look on additional information through the mixed annotations provided by broadcasting companies, sponsors of the program, and the other viewers. In the example figure, some questions the viewer may be interested will be seen. For example, an online market which sells the doll in the current movie scene, a name of actress appeared now, and a name of her hairstyle can be provided with useful external links.

The system proposed in this paper named VOAP (Video Object Annotation Platform) enables such exampled application. It consists of a systematic structure and a RDF schema which support representation and sharing of the video annotation data.

III. VIDEO OBJECT ANNOTATION PLATFORM

This section introduces a systematic structure of VOAP supporting share and the use of the video object annotation with semantically exchangeable data format.

A. *Architecturel of VOAP*

The architecture of VOAP is composed of three layers which represent three actors, namely, provider, server, and client, as shown in Figure 2.

In the provider-side, there are annotation data providers which consist of content (program) provider, economically related actors like sponsor, and ordinary viewers. In our scenario, the content provider take charge of deciding authorities for attaching annotation to a movie. Content provider distributes authorities, divided by each time seconds, to the other actors to reduce garbage data. The other actors can attach annotations to only moments of program they have an authority. Of course, there could be other way to distributes authorities, for example, Wiki [8]. However, we believe that this approach will be effective to minimize garbage annotation data and to secure rich annotation data.

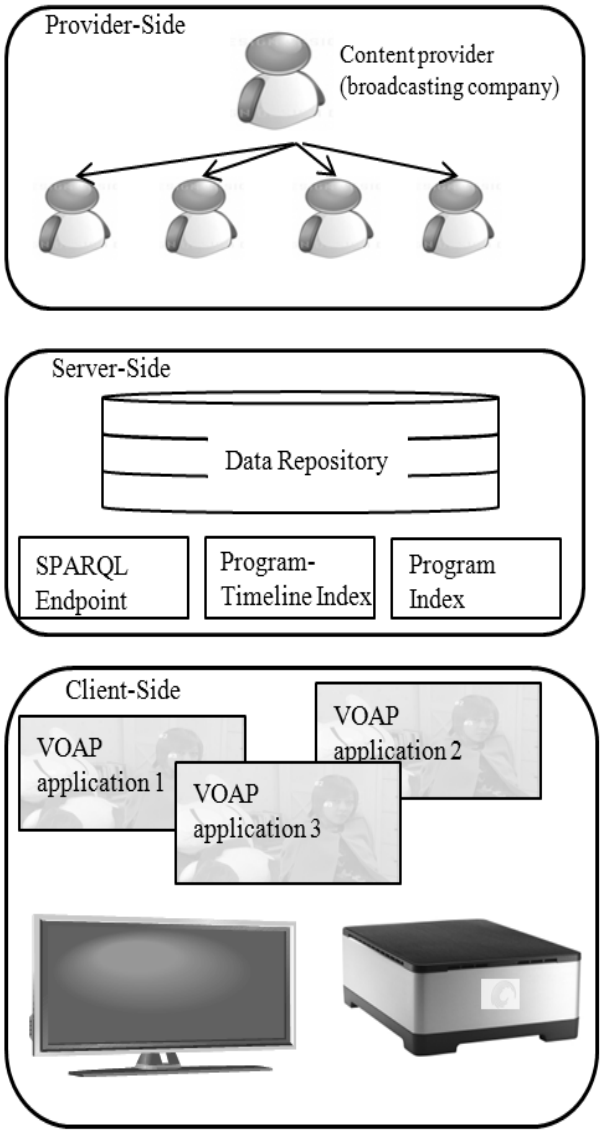


Figure 2. Artecture of VOAP

The server of the platform stores the created annotation data to be accessible to clients like applications run in the digital TVs or set-top boxes. As the annotation data are represented as RDF data formation to enhance interoperability, the server-side is equipped with a semantic data store like Jena [9]. By keeping annotation data persistent, the server has three types of access points; SPARQL, program index, program-timeline index. SPARQL provides SPARQL, as a standardized query language for semantic web data, allows a user to query on data. As proved in previous researches on benchmarks on RDF data stores, the store lacks of stable response time for large scale of data. Some of recently approaches [10] invented with massive computation power to create index data are applicable to build indexes.

The applications in clines-side provide user-oriented services with the shared annotation data as illustrated in figure1. The applications are running on the digital TV or set-top box that both are connected to the internet and are equipped with operating system for various applications.

B. Schema for Annotation Data

The video object annotation data are represented as RDF in VOAP. As a serialize-able data format, it promises interoperability and reuse of data among different applications. Figure 3 depicts a RDF data schema used for representing the annotation data. As the figure shows, the schema is related to the BBC Programs ontology [11] and FOAF (The Friend of a Friend) [12]. The first one is usable to universally publish program-related information, whereas, the last one is exploited to represent information about agent or person.

Instances of po:program from the BBC program ontology are linked as a program which have a annotation data, and foaf:person instances are linked to denote who created the annotation. Then, the annotation data which is an instance of voap:annotation has several type of attributes as followings.

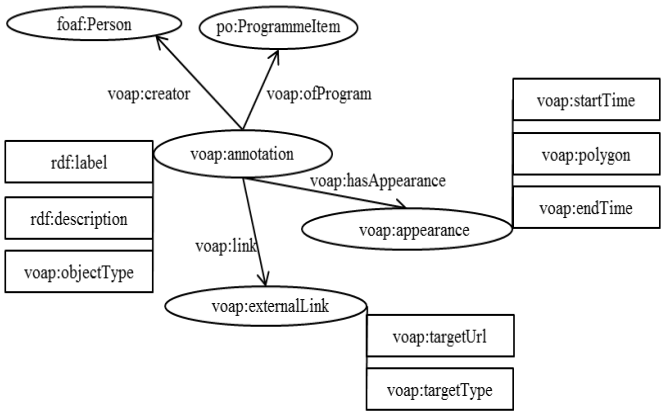


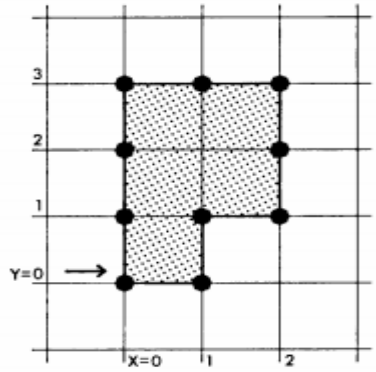
Figure 3. Ontology schema for Video-Object Annotation

TABLE I. A LIST OF ATTRIBUTES APPEARED IN FIGURE3 WHEN A CIRCLE DENOTES AN INSTANCE AND A SQUARE MEANS A ATTRIBUTE.

Attributes	**Description**
rdf:label	Label of the annotation
rdf:description	Description for the annotation
voap:objectType	A type of objet appeared and selected in the program. e.g., human, doll
Properties	**Description**
voap:creator	Indicates foaf:Person who created the annotation
voap:ofProgramm	Indicates po:ProgrammeItem the annotation is attached
voap:hasAppearance	Indicates an appearance of the annotation with start-end time and rendered polygon
voap:externalLink	Indicate an external link including explanation of the annotated object.

C. Polygon rendering

Polygon rendering is employed to designate an exact spot to be located an annotation in a scene. The computation of polygonal areas is a common operation in cartographic systems. The method of area calculation employed is dependent to some extent on the data format. A formula which was proven long before the days of computer-assisted cartography, Pick's Theorem, calculates areas of polygons whose vertices are points in a regular grid. The basic equation is AREA = GI + 1/2*GB - 1 where GI is the count of grid points inside of the polygon and GB is the count of grid points in the polygon's perimeter [13].



CODE	1	1	1	0	0	3	3	2	3	2
DX	0	0	0	1	1	0	0	-1	0	-1
DY	1	1	1	0	0	-1	-1	0	-1	0
GI	0	0	0	8	12	12	12	10	11	11

GI' = (11) - (10) = 1
AREA = (1) + 1/2*(10) - 1 = 5

Figure 4. Polygon rendering for annotating objects appeared in video

 ISBN: 978-1-61208-215-8

Using the polygon rendering algorithm in Figure 4, the area that includes an object annotated can be encoded when the annotation data is provided. Then, the area can be informed to the clients after decoding the rendered data.

IV. IMPLEMENTATION

To prove a concept of VOAP, we set up an annotation data repository at service-side and implemented a client-side application, so called, Interactive InforBox. The service-side data repository is equipped with Jena TDB [9] for serving queries on RDF data through SPARQL, RDF data query language. When he/she pause the VOD playing, the application invokes a data request query to the data repository with SPARQL constrained with a eclipsed time of current VOD and an URI (Uniform Resource Identifier) assigned to each VOD.

TABLE II. AN EXAMPLE SPARQL QUERY AND ANSWER FOR REQUESTING ANNOTATION DATA

```
SELECT ?annotations, ?apperance ?externalLink
WHERE
{
      ?annotations voap:ofProgram "http://www.etri.re.kr/voap/
programmItem/item1" .
      ?annotations voap:hasApperance ?appearance .
      ?annotations voap: externalLink ?externalLink .
      ?appearance voap:startTime ?startTime.
      ?appearance voap:endTime ?endTime.

 FILTER (?startTime >= "2012-05-28T07:18:11Z"^^xsd:dateTime
&&  ?endTime <= "2012-05-28T07:18:12Z"^^xsd:dateTime)
}
```

```
// Annotation
<rdf:Description
rdf:about="http://www.etri.re.kr/voap/annotation/annotation1">
       <rdf:label> Friends Season4 Episod 1 </rdf:label>
       <rdf:description> Monica is stung by a jellyfish and asks Joey
and Chandler to help her. The three...</rdf:description>
       <rdf:objectType>Information</rdf:objectType>
       <voap:externalLink
rdf:resource="http://www.etri.re.kr/voap/exLink/link1"/>
       <voap:hasApperance
rdf:resource="http://www.etri.re.kr/voap/apperance/apperance1" />
     </rdf:Description>

//Apperance
<rdf:Description
rdf:resource="http://www.etri.re.kr/voap/apperance/appearance1">
       <voap:stratTime>2012-05-28T07:18:09Z </voap:stratTime>
       <voap:endTime>2012-05-28T07:18:13Z </voap:stratEnd>
       <voap:polygon>CODE[1,1,1,0,0,3,3,2,1,7]DX[0,0,0,0,0,1,-1,-
1,1,7,]DY[0,2,0,4,0,1,-2,-1,1,7]GI[1,2,0,12,0,10,24,11,13,17]
</voap:polygon>
</rdf:Description>

//External Link
<rdf:Description
rdf:resource="http://www.etri.re.kr/voap/exLink/link1">
       <voap:targetUrl rdf:resource
="http://en.wikipedia.org/wiki/Jennifer_Aniston"/>
       <voap:targetType>WebLink</voap:targetType>
</rdf:Description>
```

Upon requested, it returns a video object annotation data to be shown over the current scene. The answer for the requested SPARQL includes a set of annotations with appearance of the annotations and external links related to the featured objects as shown in Table 2.

The Interactive InforBox application reads the annotation data and decodes rendered polygons to pop-up InforBox window. As Figure 5 shows, a television view can see the colored buttons with corresponding objects over the current scene. When the button clicked by the viewer, the screen is redirected to a web page containing related information of the object, so that, the viewer can supplied with additional information, such as, details on the object, commercial advertisements, and other related contents .

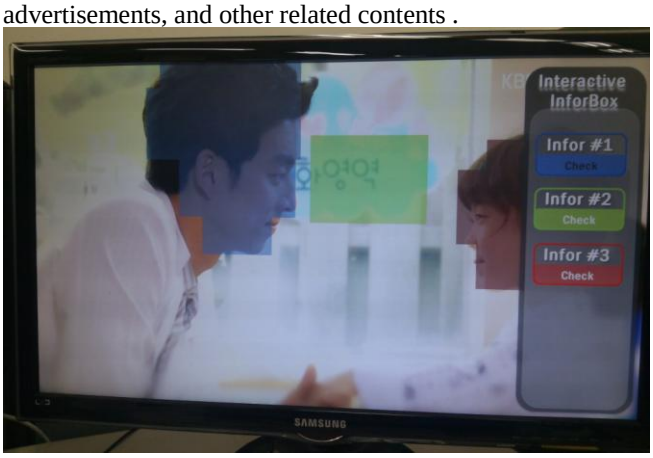


Figure 5. Imange of Intractive InforBox implementation

V. CONCLUSION

This paper introduced the VOAP supporting share and use of the video object annotation. The platform is consisted of a data model for encoding video object annotation data and the data exchange scenario. RDF based data model promises an information exchanging environment universally usable in heterogeneous interactive TV platforms. And, the data exchanging environment can open a business model by adding a new information channel to the television. As a work in progress, we implemented a user-side application proving the proposed concept.

By referring web editorial tools like Wiki, a strategy to resolve multiple annotation attached to the same object at the same moment need to be developed . Further developments on the platform will cover the applications useable in the platform for enhancing the information transfer among consumer, advertisers (company) and mediation. An annotation tool equipped with object tracing technology will be also provided to ensure easier annotating on the moving video object.

ACKNOWLEDGMENT

This work was supported by IT R&D Program of MKE/KEIT/KCC, [Development of Social TV Service Enabler based on Next Generation IPTV Infrastructure].

 ISBN: 978-1-61208-215-8

REFERENCES

[1] Y. Xiao, X. Du, J. Zhang, F. Hu, and S. Guizani, Internet Protocol Television (IPTV): The Killer Application for the Next-Generation Internet, Communications Magazine, IEEE , vol. 45, no. 11, pp. 126-134, November 2007.

[2] B. Schopman, D. Brickley, L. Aroyo, C. V. Aart, V. Buser, R. Siebes, L. Nixon, L. Miller, V. Malaise, M. Minno, M. Mostarda, D. Palmisano, and Y. Raimond, NoTube: making the Web part of personalised TV, In Proceedings of the WebSci10: Extending the Frontiers of Society Online Poster Track. Apr 2010.

[3] H. Jeon, Y. Shin, M. Choi, J. Rho, and M. Kim, User Adoption Model under Service Competitive Market Structure for Next-Generation Media Services, ETRI Journal, vol.33, no. 1, Feb. 2011, pp. 110-120.

[4] Wikipedia, Telestrator, Available at http://en.wikipedia.org /w/index.php?title=Telestrator&oldid=180785495, 2006, [retrieved: May 2012].

[5] J. Vendrig and M. Worring, "Interactive adaptive movie annotation", IEEE Multimedia, vol 10, issue 3, pp. 33-37, IEEE Computer Society.

[6] D.B. Goldman, C. Gonterman, B. Curless, D. Salesin, and S. M. Seitz, Video object annotation, navigation, and composition, In Proceedings of the 21st annual ACM symposium on User interface software and technology (UIST '08). ACM, New York, NY, USA, pp. 3-12.

[7] Wikipedia The Free Encycropedia, Available at http://wikipedia.org, [retrieved: May 2012].

[8] B. Leuf and W. Cunningham. The Wiki Way: Quick Collaboration on the Web. Addison-Wesley Professional, April 2001.

[9] B. McBride, Jena: a semantic Web toolkit, Internet Computing, IEEE, vol.6, no.6, pp. 55-59, Nov/Dec 2002.

[10] S. Yun, M. Song, H. Namgung, S. Yang, H. Kumar, and H. Kim, Improving the workflow of semantic portals using M/R in cloud platform, In proceedings of KEOD 2010-International Conference on Knowledge Engineering and Ontology Development, 2010, pp. 485-488.

[11] BBC Programmes Ontology, Avalable at http://www. bbc.co.uk/ontologies/programmes/2009-09-07.shtml, [retrieved: May 2012].

[12] M. Graves, A. Constabaris, and D. Brickley, FOAF: Connecting People on the semantic Web. Cataloging & Classification Quarterly, 43(3/4), 2007, pp. 191-202.

[13] H.S.M. Coxeter, Introduction to Geometry, pp. 208-210, John Wiley and Sons, Inc., New York, 1961.

 ISBN: 978-1-61208-215-8