

Design of Scene Knowledge base system based on Domain ontology

Wonjoo PARK*, Minhoo Han*, Jeong-Woo Son*, Sun-Joong Kim*

*Smart Media Research Department, Hyper-Connection Media Research Laboratory,

ETRI(Electronics and Telecommunications Research Institute), Korea

[wjpark, mhhan, jwson, kimsj]@etri.re.kr

Abstract— In light of the rapid growth VoD clips commerce services, there are more needs video and scene metadata to be searchable and re-usable. We propose a design of scene knowledge base system based on a domain knowledge ontology. The domain knowledge ontology is knowledge base to describe some scenes which who, when, where, what and how. It has actors, objects, times, locations, activities, moods, brands, and so on. It also pursues to include as many real and virtual world events as possible. After all, scene knowledge base are scene metadata and some events ontology based on domain ontology. They can be used for VoD clips, media commerce, marketing, advertisement services.

Keywords— scene metadata, scene knowledge base, scene ontology, domain ontology, domain knowledge base

I. INTRODUCTION

Nowadays, VoD clips services are very popular and is also expected to increase further in the next few years. Already, the YouTube and Naver, Kakao clip service market is very active and the clips market is expected to become more intense as the latest Facebook followers join the market. So, most of contents providers like movie director, broadcasting company want to be searchable and reusable their own clips and eventually to make some profits. Therefore, clips need to be tagged events, peoples, locations, moods, and so on like some metadata. However, most of metadata generation and tagging services have been focused on movies, dramas, news based on whole contents. It is not adjustable to clip service which needs short time length and a few events. To tackle this problem, we propose a design of scene knowledge base system based on domain knowledge ontology.

II. RELATED WORKS

A scene knowledge base is a scene metadata in the narrow sense, and the description of scene in a broad sense. To establish a good scene knowledge base system of real world or virtual world is very difficult and almost impossible task. However, it is critical to have scene knowledge base in order to search and re-use the scene. Also when the scene knowledge system has ontology structure, it has some

advantages to share common understanding of the structure of information, to enable reuse of knowledge, etc [1].

G. Ye, Yitong Li et al. propose EventNet. It is event concept ontology and a mid-level representation of complex events in video. They pursue a large scale structured concept library complex event detection in video. They adopt the CNN architecture as deep learning model to perform deep feature form the video content and use event ontology to cover as many real world events and concept as possible. Also, they use wikiHow which is an online forum to gather human daily life[2][3].

S. Atkinson et al. propose the ontology and metadata design for digital film production assets. The deep film access project (DEAP) aims to unlock latent opportunities that exist within the big and complex data sets generated by industrial digital film production. They focus on developing methodology for the integration of the data and metadata generated through film production, by capturing, archiving and making accessible the diverse range and levels of expertise which exist within film-making[4].

Kadar et al. propose SceneNet which is a perceptual ontology for scene understanding. They try to establish online perceptual ontology database for scene understanding[5].

J. Hartz et al. propose knowledge base of ontological concepts for high-level scene interpretation. Their approach focus on a description language for spatial object arrangements, applies the learning procedure to a concrete real world domain, evaluate the results in an operational scene interpretation system. They adopt ontological concept to learn and find object in a space[6].

J. Kim et al. propose emotion-media metadata authoring tool, emotion-media transmission server and terminal. In addition, they developed scent display device, an olfactory media broadcasting system based on MPEG-V for emotion augmentation of video services. The emotion-media authoring tool synchronizes the scent effect metadata with the TV program, based on MPEG-V[7].

III. SYSTEM OVERVIEW

A scene is very similar with clip and our system uses the term, scene. A Scene is a short time length video clip and a part of whole video file and has a few events.

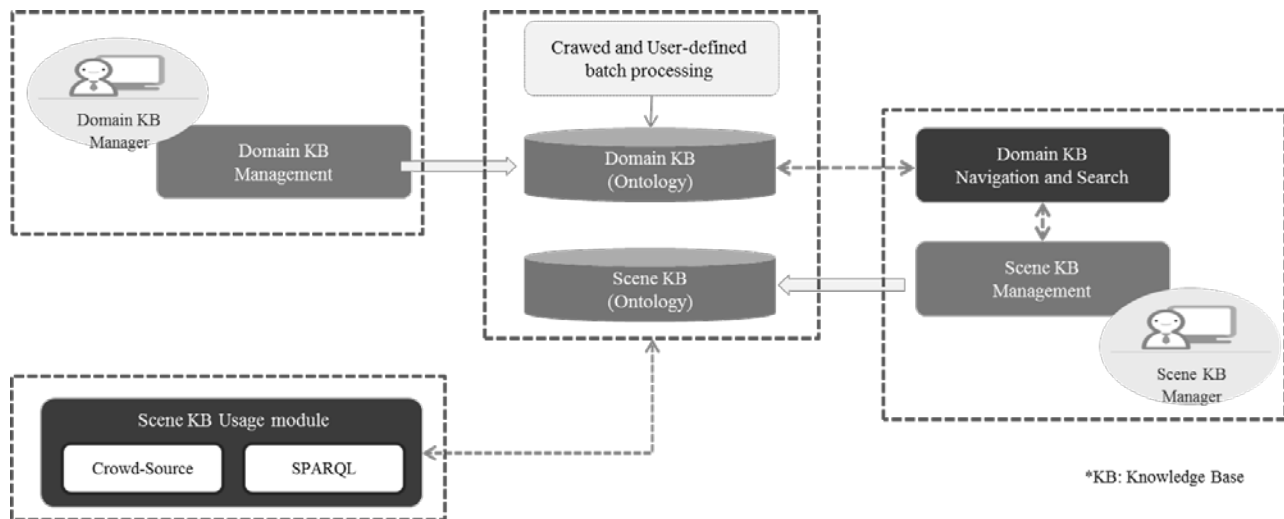


Figure 1. Scene Knowledge base system architecture

Figure 1 shows the overall structure of the proposed system. As shown in this figure, the proposed system is composed of mainly two knowledge bases and three modules: a Domain Knowledge Base, a Scene Knowledge Base, a Domain Knowledge Base Management module, a Scene Knowledge Base Management module and Scene Knowledge Base Usage module. To establish scene knowledge base, we use a domain knowledge base. The domain knowledge base pursues to include as many real world and virtual world events and human life activities as possible. First, we define the domain knowledge base which is based on the domain ontology. To be used as scene metadata, it has movie actors, animals, activities, time, locations, objects, and so on. When the domain knowledge base management module defines some domain ontology, it reuses existing ontologies like other movie ontology, korea actor ontology, DBpedia, etc. Also, it transforms crawled data with open-API and user-defined data to RDF and adds instances. And then a Scene knowledge base manager searches or navigates domain ontology and tags scene metadata while watching the scene.

Figure 2 shows a Crowd-Source system example to establish scene knowledge base. A Scene knowledge base manager uploads clip file and file information. In this step, it may use domain ontology to find content title, main actors. And then the manager plays and watches the scene. To tag metadata who showed up, where is the place, what kind of activity, he navigates or searches the domain ontology like an actor, location, activities. Also, it wants to tag time, emotions, moods, clothes, landmarks, and so on. It eventually has 4W 1H property - who, when, where, what, how - about any event. It can tag multiple events and objects to the scene.

Figure 3 shows a scene ontology example. The specific scene URI has start time, end time, activities, actors, locations, time, and video titles. As most of values have own URI and label, they can be searched by keyword or SPARQL query.



Figure 2. Scene KB Usage module by Crowd-Source

http://etri.rdf.kr/resource/Cinefox_2016-00-02-11_01

Property	Value
etrip:comeoutAction	- etrip:comeoutActionSeparating - etrip:comeoutActionWalking
etrip:comeoutActionLabel	- 갈등 - 회개
etrip:comeoutEmotion	etrip:comeoutEmotionSorrow
etrip:comeoutEmotionLabel	슬픔
etrip:comeoutMovieActor	- etrip:comeoutMovieActorParkMinyoung - etrip:comeoutMovieActorParkMinyoung
etrip:comeoutMovieActorLabel	- 박민영 - 박민영
etrip:comeoutObjectClothing	etrip:comeoutObjectClothingCoat
etrip:comeoutObjectClothingLabel	코트
etrip:comeoutPlace	etrip:comeoutPlaceStreet
etrip:comeoutPlaceLabel	길거리
etrip:comeoutTimePeriod	etrip:comeoutTimePeriodMidnight
etrip:comeoutTimePeriodLabel	심야
etrip:endTime	189
etrip:filePath	etrip:filePathHomeRepositoryCinefox_2016-00-02-11.mp4
etrip:startTime	182
etrip:type	etrip:DevideSceneKnowledge
etrip:userDevideVideo	http://etri.rdf.kr/resource/Cinefox_2016-00-02-11
etrip:video	http://etri.rdf.kr/resource/20060576
etrip:videoName	그날의 눈이

Figure 3. A Scene ontology example

The scene knowledge base usage module generates scene metadata using scene ontology. It will be expanded to a structure that generates rich metadata by using text mining and object recognition technology.

IV. CONCLUSIONS

We proposed a design of scene knowledge base system based on a domain knowledge ontology. The domain knowledge ontology is knowledge base to describe some scenes which who, when, where, what and how. It has actors, objects, times, locations, activities, moods, and so on. It also pursues to include as many real and virtual world events as possible. After all, scene knowledge base are scene metadata and some events ontology based on domain ontology. We are developing this system and plan to research automatic metadata generation based on text mining[8] and object recognition technology.

ACKNOWLEDGMENT

This work was supported by Electronics and Telecommunications Research Institute (ETRI) grant funded by the Korean government. [16ZC2200, Development of programmable interactive media creation service platform based on open scenario].

REFERENCES

- [1] NOY, Natalya F., et al. Ontology development 101: A guide to creating your first ontology. 2001.
- [2] Ye, G., Li, Y., Xu, H., Liu, D., & Chang, S. -. (2015). EventNet: A large scale structured concept library for complex event detection in video. Paper presented at the MM 2015 - Proceedings of the 2015 ACM Multimedia Conference, pp. 471-480.
- [3] <http://www.wikihow.com/Main-Page>
- [4] S. Atkinson, J. Lehmann and R. Evans, "The Deep Film Access Project: Ontology and metadata design for digital film production assets," Big Data (Big Data), 2014 IEEE International Conference on, Washington, DC, 2014, pp. 1-4.
- [5] Kadar, I., & Ben-Shahar, O. SceneNet: A perceptual ontology for scene understanding, 13th European Conference on Computer Vision, ECCV 2014, Zurich, Switzerland, Volume 8926, 2015, pp. 385-400.
- [6] J. Hartz and B. Neumann, "Learning a knowledge base of ontological concepts for high-level scene interpretation," Machine Learning and Applications, 2007. ICMLA 2007. Sixth International Conference on, Cincinnati, OH, 2007, pp. 436-443.
- [7] Jeong Do Kim, Ji Hoon Choi, Seung Ju Lim, Sung Dae Park, Jung Ju Kim, and Chung Hyun Ahn, "Development of Scent Display and Its Authoring Tool," ETRI Journal, vol. 37, no. 1, Feb. 2015, pp. 88-96.
- [8] J. Son, H. Yoon, K. Cho, W. Ryu, "Consolidation of Subtasks for Target Task in Pipelined NLP Model, ETRI Journal, Volume 36, Number 5, 2014 pp. 704-713.



Wonjoo Park received her MS degrees in information and communication engineering from Chungnam National University, Daejeon, Rep. of Korea in 2000. She joined ETRI, Rep. of Korea in 2000, where she is currently senior researcher. Her research interests include data mining, topic model, and ontology.



Minho Han received her BS and MS degree in computer engineering from Chungnam National University, Daejeon, Rep. of Korea in 2001. He joined ETRI, Rep. of Korea in 2000, where he is currently senior researcher. His research interests include natural language processing and information retrieval.



Jeong-Woo Son received his MS and Ph.D. degrees in computer engineering from Kyungpook National University, Daegu, Rep. of Korea in 2007 and 2012 respectively. Since 2013, he has been with ETRI, Daejeon, Rep. of Korea. He focuses on machine learning, NLP, and information retrieval.



Sun-Joong Kim received her BS degree in computational statistics and her MS degree in computer science from Chungnam National University, Daejeon, Rep. of Korea, in 1989 and 2000 respectively. In February 1989, she joined ETRI, Daejeon, Rep. of Korea, where she is currently principal researcher and director. Her research interests include convergence service control, smart TV, content knowledge mining.