International Journal of Internet, Broadcasting and Communication Vol.11 No.3 71-76 (2019)
http://dx.doi.org/10.7236/IJIBC.2019.11.3.71

IJIBC 19-3-10

One-Click Marketing Solution for Mobile Videos

Jae Seung Lee[††], Seung Heon Lee[††], Jin Woo Jang[†], Hyun Bin Kim[†], Ga Young Nam[†], Suk Ho Lee[†]

[††]*SuperSell Corporation*
[†]*Dept Software Engineering, Dongseo University*
petrasuk@gmail.com

Abstract

In this paper, we propose a simple one-click marketing solution for mobile devices which can advertise a product which is embedded in a mobile video while watching the video on a smartphone. If a specific product of interest appears in the video to the user, one can simply click on the product in the video and a pop-up window with information about the product is proposed. The implementation of the system is expected to enable users to gain real-time information about the product while watching the video without having to search for the product again after watching the movie, and thereby facilitating more mobile commerce. We use a two-fold system to prevent the failure of tracking which often occurs on a single online tracking system, so that the user cannot always get the commercial product information.

Keywords: *Mobile Advertisement, Object Tracking, One Click, Mobile Video, Smart-phone*

1. Introduction

Recently, the ubiquitous screen era has arrived where video can be viewed anytime and anywhere due to the widespread use of smart phones. In 2017, the worldwide traffic for mobile videos which is estimated to be 1.4 Hexabytes is expected to grow six-fold to 8 Hexabytes in 2019. Data traffic for mobile videos in South Korea is also expected to increase 5.8 times in 2021 compared to 2016. Along with the expansion of video streaming platforms and the expansion of mobile video contents, the mobile video advertisement market is also growing rapidly. When watching TV dramas and online wed videos, viewers often wonder what the products are that the actors/actresses are wearing. Accordingly, viewers are going through a research on the Internet to obtain the product information, and, without knowing the exact product name, names such as "My golden colored Life, coat" become more popular than the original names(see Figure 1). This is due to the fact, that these products become famous on the Internet making the search word related to the movie more popular than the original true product name. However, it takes some time that these search words become popular, and the viewers cannot obtain research results with these names directly after watching the movie and have to go through an online re-search to get information about the product. In

Manuscript Received: June. 27, 2019 / Revised: July. 5, 2019 / Accepted: July. 12, 2019
Corresponding Author: petrasuk@gmail.com
Tel: +82-051-320-1744, Fax: +82-051-327-8955
Division of Computer Engineering, Dongseo University, Korea

addition, in the case of modern dramas and videos, information is often missing due to the difficulty of searching for such products.

Figure 1. Example of search for popular product.

Therefore, in this paper, we propose a system which shows information of the product in the movie by just clicking on the product while watching the online video. Figure 2 shows the concept of this service. When there is a product in which the viewer is interested the viewer clicks on the product (the coat in Figure 2) and a pop-up window with information about the product and a link to the purchase site is opened, enabling a real-time purchase. The viewer saves the time for searching on the Internet, and immediately can buy the product. This kind of advertising will also increase the sales volume as the time for changing the desire for purchase is reduced.

Figure 2. Concept of the one-click e-commerce service.

2. Overall System Configuration

Figure 3 illustrates the overall system configuration for realizing the service described in the introduction. As the actor (and so the products the actor is wearing) moves constantly in the video, the products have to be tracked in the movie so that the product can be clicked on anytime to obtain the information of the product.

However, it is well-known that existing object tracking algorithms are liable to fail in the tracking no matter how robust the tracking algorithm is. If the tracking is doing online, there surely will appear cases in which the tracking fails, and the viewer, while clicking on the object, may get no information about the product. This failure in the tracking will result in the withdrawal of the service users. Therefore, the react of the service should be 100% successful. This kind of success rate cannot be obtained with online tracking. Therefore, we propose a two-fold system. The first system is an offline interactive tracking system, which tracks the object of interest with the help of the interaction of the video advertisement maker.

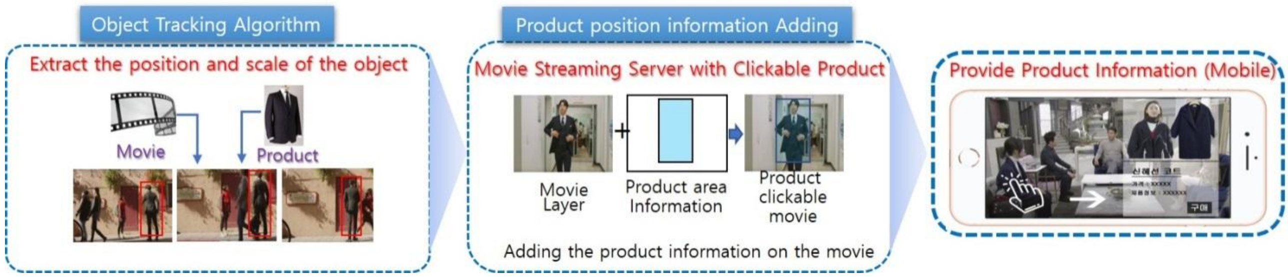

Figure 3. Overall System Diagram.

The video advertisement maker chooses manually the tracking object in the first frame. The information of the tracking object is then stored in a temporary memory and used for the tracking of the object. If the tracking fails, the object is manually retargeted by the video advertisement maker. The location and the size of the product object is then stored into a database. As the tracking is done offline with the interaction of the video advertisement maker, the location and size information of the product is 100% correct.

The information obtained by the offline tracking system is then forwarded to the online movie streaming system. The online movie streaming system relates the video with the product area information and makes the video clickable. So, when the video containing these product domain information is serviced via a server to the mobile device, the product will be clickable in the mobile video. Using this two-fold system, the viewer can have always a response to the click on the product.

3. Deep learning based Object Tracker for the Offline Advertisement Embedding System

We utilize a Siamese artificial neural network-based deep learning tracking algorithm [1, 2] together with some online information as the object tracking algorithm. The Siamese artificial neural network consists of two general neural networks, where the inputs of the networks are the two different images for which we want to compare the similarity. The output of the networks are then inputted into a similarity metric, which has a large value if the two images are similar, while has a small value if the images are different. In the SiamRPN tracking method, this kind of similarity is used to track the object by choosing the next position of the object as the one which has the largest similarity value with the object region in the previous frame. In this paper, we use a modification of the SiamRPN which considers also the online information together with the region proposal of the SiamRPN[3,4], i.e., we use the following matching score:

$$matching\ \ score\ = \sum_i \|\boldsymbol{h}(p_t^{n-1}) - \boldsymbol{h}(p_i^n)\|\, D(p_t^{n-1}, p_i^n) \quad (1)$$

where $\boldsymbol{h}(p_k^{n-1})$ and $\boldsymbol{h}(p_i^n)$ are the histogram vectors at the positions p_k^{n-1} (the position of the target in the previous frame) and p_i^n (the position of a target candidate position in the current frame), respectively, and $D(p_t^{n-1}, p_i^n)$ is the distance between p_k^{n-1} and p_i^n. For the distance function we use the square of the L2 norm difference norm:

$$D(a, b) = \|a - b\|^2 \quad (2)$$

Therefore, the matching score is large if the target in the previous frame and the target candidate in the current frame have similar colors and are close to each other. By choosing from the regions proposed by the

SiamRPN the one having the largest matching score, the next target region can be obtained more robustly.

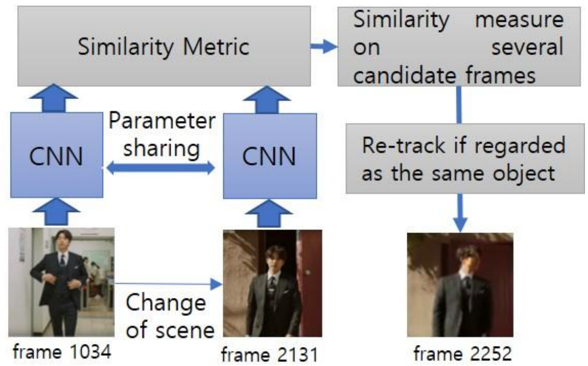

Figure 4. Siamese network based tracking framework.

4. Method of Processing the Touch Event

Figure 5 shows the diagram of handling the event of touching the product in the video. When the user clicks the product in the image, the frame number is calculated by dividing the time of the click by the total time of the video. This current frame number and the position of the click information are sent to a pre-saved database as a query for the matching of a product region. The information are the result of the applying the offline object tracking algorithm to a specific product. Each record has XMIN, YMIN, XMAX, and YMAX values which correspond to the X, Y coordinates of the left top and the right bottom corner of the product area, respectively. These coordinate values are sent to the front-end as a return from the database. At the front end, the (X,Y) coordinate values of the clicked position is evaluated if it lies in a box region which has (XMIN,YMIN) as the top left corner, and the (XMAX,YMAX) coordinate as the bottom right corner. If it is lies within this area, a query is sent to the product information database, and the information for this product is displayed on the screen through a pop-up window. If the click position is not within the area, the video continues to being played normally without interruption. Figure 6 shows the whole diagram of the touch event handling process

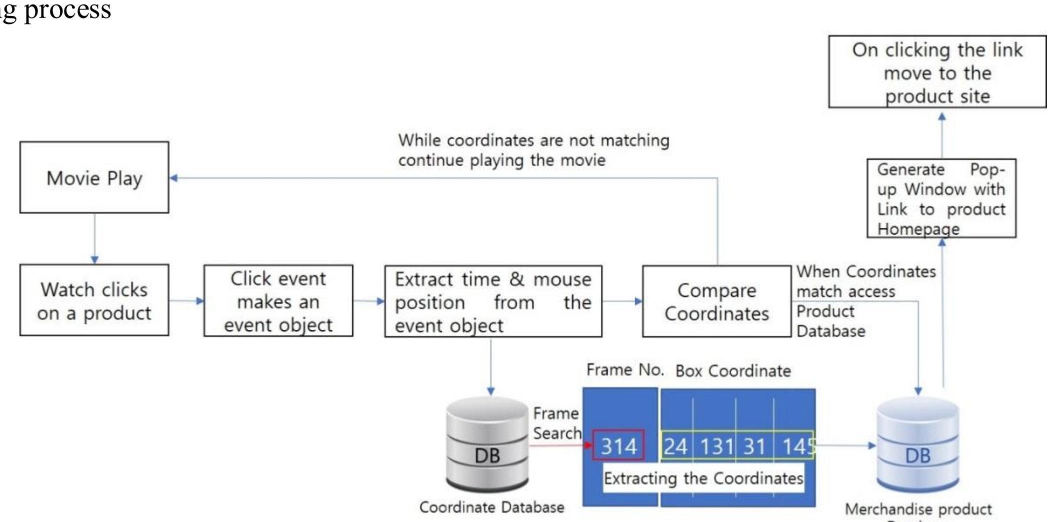

Figure 5. Diagram of the touch event handling process.

5. Demonstration of the System

We implemented the system with node.js for the server side and JavaScript and bootstrap for the front-end side. Figure 6 illustrates the how the program will run if it is run on the server side, i.e., on a desktop PC. In this case, the mouse click is regarded as the touch event and the pop-up window appears as a part of the full screen playing window.

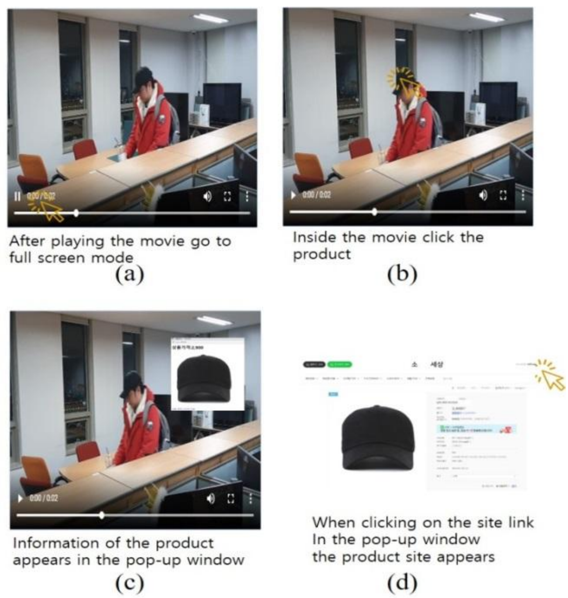

Figure 6. Illustration of running the system on the server side(PC).

Figure 7 shows captured photos of a real demonstration of the system when run on a mobile device. The real size of the full screen depends on the mobile device. Therefore, the coordinates of the mobile screen have to be computed relatively to those on the server. Therefore, we first gain the full screen size of the mobile device, and then re-compute the coordinates according to the ratio of the sizes of the PC screen and the mobile phone screen.

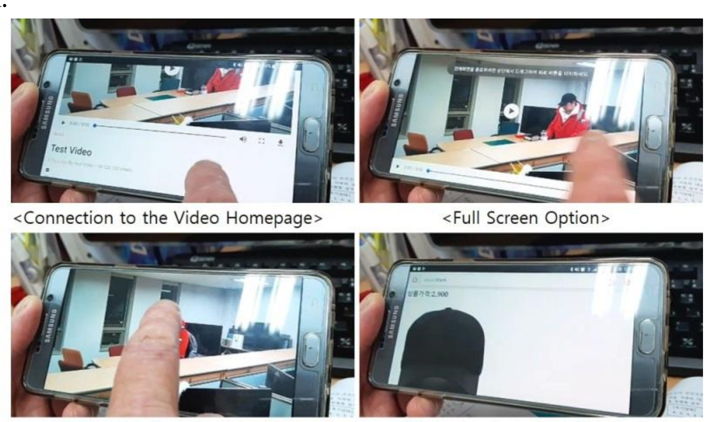

Figure 7. Real-time demonstration on running the system on a mobile phone connected to the server via the internet.

5. Conclusion

In this paper, we proposed a one-click advertisement system that enables a real-time purchase at the viewing time of a mobile video. We proposed a two-fold system, where the tracking of the product is being performed by a stand-alone type software with the interaction of the advertising video maker to correct for the failures of the tracking, and the real-time play of the video with the information of the product position is performed via a client-server system, which gets the result of the tracking from the database. We showed in the experiments how the system functions. To reduce the amount of labor of interaction to get the correct coordinates of the product's position, it is required to use a good tracking algorithm, which is one of the topics of further studies for the proposed system.

Acknowledgement

This work was supported by the Technology development Program(S2644388)funded by the Ministry of SMEs and Startups(MSS, Korea)

References

[1] H. Park and S. Lee, "User Selection Based Backpropagation for Siamese Neural Networks in Visual Filters," *Lecture Notes in Electrical Engineering*, Vol. 488, pp 109-115, 2016..

[2] Y. Wu, J. Lim, and M. Yang, "Object tracking benchmark," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 37, No. 9, pp. 1834–1848, Sep. 2015. DOI: 10.1109/TPAMI.2014.2388226

[3] B. Li, J. Yan, W. Wu, Z. Zhu, and X. Hu, "High performance visual tracking with siamese region proposal network" *In: 2018 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 8971–8980, June 2018. DOI: 10.1109/CVPR.2018.00935

[4] Z. Zhu, Q. Wang, B. Li, W. Wu, J. Yan, and W. Hu, "Distractor-aware Siamese Networks for Visual Object Tracking," *European Conference on Computer Vision (ECCV)*, pp.103-119, Sep. 8-14, 2018.

[5] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," In: *2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)*, pp. 779–788, June, 2016. DOI: 10.1109/CVPR.2016.91

[6] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *Advances in Neural Information Processing Systems*, Vol. 28, pp. 91–99, 2015.

[7] H. Nam and B. Han, "Learning multi-domain convolutional neural networks for visual tracking," In: *2016 IEEE Conference on Computer Vision and Pattern Recognition(CVPR),* pp. 4293–4302, June, 2016. DOI: 10.1109/CVPR.2016.465

[8] L. Bertinetto, J. Valmadre, S. Golodetz, O. Miksik, and P.H.S. Torr, "Staple: Complementary learners for real-time tracking," In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. pp. 1401–1409, June, 2016. DOI:10.1109/CVPR.2016.156

[9] L. Bertinetto, , J. Valmadre, J.F. Henriques, A. Vedaldi, and P.H.S Torr, "Fully-convolutional siamese networks for object tracking," In: *ECCV 2016 Workshops,* pp. 850–865, 2016.

[10] G. Ning, Z. Zhang, C. Huang, X. Ren, H. Wang, C. Cai, and Z. He, "Spatially supervised recurrent convolutional neural networks for visual object tracking," In: *2017 IEEE International Symposium on Circuits and Systems (ISCAS)*, pp. 1–4, May, 2017. DOI:10.1109/ISCAS.2017.8050867