Demo: Panoramic Streaming using Named Tiles

Kazuaki Ueda
KDDI Research, Inc.
kz-ueda@kddi-research.jp

Yuma Ishigaki
Osaka University
y-ishigaki@ist.osaka-u.ac.jp

Atsushi Tagami
KDDI Research, Inc.
tagami@kddi-research.jp

Toru Hasegawa
Osaka University
t-hasegawa@ist.osaka-u.ac.jp

ABSTRACT

This demonstration shows an efficient panoramic streaming application with ICN functionalities. The camera device splits its field-of-view into multiple named tiles, and clients request minimum tiles for their field-of-view. These named tiles are cached at the intermediate routers and the camera device can reduce the amount of traffic on its access network. By utilizing the ICN's merits, this application can work with limited resources, like IoT environment.

CCS CONCEPTS

• **Networks** → *Network experimentation*; • **Information systems** → Multimedia streaming;

KEYWORDS

Information Centric Networking, Panoramic Video

ACM Reference format:
Kazuaki Ueda, Yuma Ishigaki, Atsushi Tagami, and Toru Hasegawa. 2017. Demo: Panoramic Streaming using Named Tiles. In *Proceedings of ICN '17, Berlin, Germany, September 26–28, 2017,* 2 pages.
https://doi.org/10.1145/3125719.3132093

1 INTRODUCTION

360 degree panoramic video is becoming popular due to amusingness, future outlook and adaptive flexibility. Its key feature is that users can navigate the video to areas in which they are interested by high-spatial resolution. This interactivity expanses the domain of its applicability, like a surveillance camera, an equipment inspection, or a live streaming for sports. Omnidirectional cameras which can create panoramic video are also becoming popular and developed by many venders [6]. For the further improvement of its applicability, working in the resource constrained environment, i.e., limited computer and network resources, is an important issue.

Usually, the camera device has only limited computer resources for encoding with high compression rate. And the network resource is also inadequate to transmit raw data of high-spatial resolution video streaming. To operate under this environment, the efficient panoramic streaming is required. This demonstration shows an efficient panoramic video streaming using named tiles on ICN.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
ICN '17, September 26–28, 2017, Berlin, Germany
© 2017 Association for Computing Machinery.
ACM ISBN 978-1-4503-5122-5/17/09…$15.00
https://doi.org/10.1145/3125719.3132093

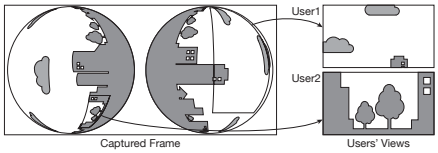


Figure 1: A sample of captured frame and user's view

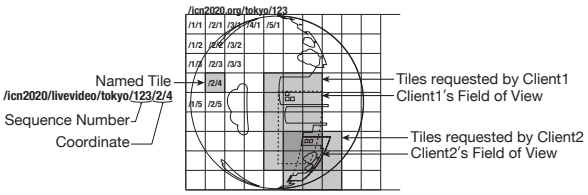


Figure 2: Captured frame is split into multiple named tile. Clients request necessary tiles for their field-of-view.

2 TILE-BASED PANORAMIC STREAMING

In panoramic video, users see only a part of the whole captured frame. Figure 1 shows the relationship between a captured frame and a user's view. To reduce the bandwidth usage, the server can encode only the requested area for each user, but it causes heavy computer processing. To avoid this on-the-fly video encoding, tile-based streaming technique is well known [1, 7]. The captured video is subdivided into non-overlapping tiles and each tile is independently compressed using a video encoder, e.g., H.264/AVC, H.265/HEVC or Motion JPEG. Clients request only necessary multiple tiles according to their field-of-view. This technique can lower the complexity of the video encoding and processing delay.

As tile-based streaming is based on the retrieval of multiple content objects, this technique can be efficiently combined to ICN. Figure 2 shows an example of the *named tile-based panoramic streaming*. Each tile has a unique name which consists of a time sequence number and a coordinate, e.g. /icn2020/livevideo/tokyo/123/2/4. And clients issue interests for the minimum necessary tiles covering their field-of-view. Since each tile is a named object of ICN, it can be cached at intermediate routers and natively delivered in a multicast manner. Thus even if there are many clients, the camera device only transfers up to the whole tiles at once. Named tile-based panoramic

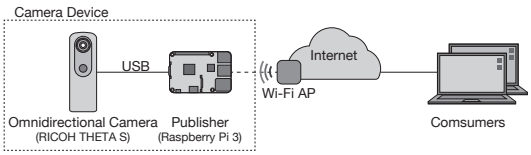
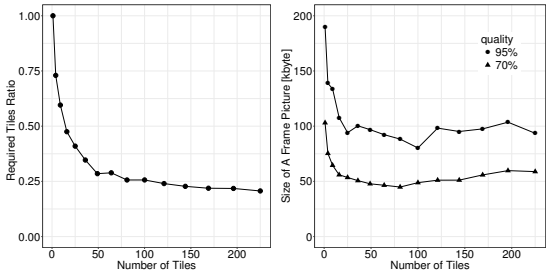


Figure 3: Overall of our system consisted of a omnidirectional camera, publisher and consumers



(a) Ratio of Required Tiles (b) Downloaded Data Size

Figure 4: Relationship between the number of tiles and required tiles/frame size

streaming can reduce redundant process and forwarding, thus it can work with resource constrained environment.

3 IMPLEMENTATION

We implement the named tile-based panoramic streaming on NDN [3]. Figure 3 shows the overall of our system. The camera device consists of an omnidirectional camera [6] and a tiny computer, Raspberry pi 3 [5]. The camera provides a dual-fisheye video stream to the computer which works as a NDN publisher. The computer slices and encodes the provided video frames using OpenCV [4]. The application uses JPEG/Motion JPEG as the encoding format and each tile is a small JPEG image which consists of the frame.

A client first obtains a metadata from publisher, that has frame definition, frame rate, tile size and current time sequence number. The metadata is used to calculate minimum necessary tiles for covering user's field-of-view that is controlled from user's controller with pan-tilt-zoom operation. The client sends interest packets for the necessary tiles and makes a patchwork image from retrieved tiles. The image is converted from a dual-fisheye image into a plane image using OpenGL [2]. Thus the user can watch the panoramic streaming without retrieving a whole frame. The time sequence number increases in accordance with the frame rate determined by the camera's specifications and the publisher's performance. The consumer calculates the current time sequence number from the frame rate and the elapsed time, and requests the tiles tied with the number. For the synchronization, the consumer acquires the metadata by sending interest every five seconds with "MustBeFresh" option. If a consumer cannot reflect a new video frame within frame interval due to network conditions and/or processing delays, it skips over some sequence. This works as a frame rate adaptation, and no need for preparing tiles with different bit rate is advantage for improving ICN efficiency, e.g. cache hit ratio. The key point of this implementation is that the camera device processes only the two-dimensional image processing and clients are responsible for the heavy, e.g., three-dimensional, processing.

In this demonstration, we put one camera device (NDN publisher) in a demonstration room, and use a simple topology that one notebook (NDN consumer) directly connects to the camera device via Wi-Fi router. We can see the panoramic streaming and also control its field-of-view by using gamepad.

4 EVALUATION

The tile size greatly affects the performance of our system. First, tile size affects the overhead of constructing user's view. As shown in the Figure 2, area of tiles should be larger than user's field-of-view, and large tiles tend to hold redundant part because of their coarse grained representation. Figure 4 (a) shows the relationship between the ratio of required tiles to the whole tiles and the total number of tiles which corresponds to tile size. The required tiles depend on user's field-of-view, and this graph plots an average number of required tiles in 25 randomly selected field-of-views. When the number of tiles increases, this means the smaller tile size, the ratio of required tiles gets smaller because it can represent user's field-of-view precisely. However, the small tile size lowers the compression ratio, and configuring too small size increases the total data volume for transmitting a frame. Figure 4 (b) illustrates the download data size under two JPEG compression qualities. According to the tendency of this result, 50 to 100 tiles is possible to reduce traffic volume by half. The system performance includes transferring data size depends on not only the tile size but also tile shape and projection format [1]. But these aspects are part of the future work and lie outside the scope of this paper.

5 CONCLUSION

This paper describes the effectiveness of combination of tile-based panoramic streaming and ICN functionalities. Our demonstration shows that our panoramic streaming can work with poor computing and network resources. The remaining work is to evaluate the behavior of our video streaming in the realistic wide area network.

ACKNOWLEDGEMENTS

The work for this paper was performed in the context of the Horizon2020/NICT EU-JAPAN ICN2020 project Grant Agreement No. 723014 and Contract No. 184.

REFERENCES

[1] Xavier Corbillon, Alisa Devlic, Gwendal Simon, and Jacob Chakareski. 2016. Viewport-Adaptive Navigable 360-Degree Video Delivery. *CoRR* abs/1609.08042 (2016). http://arxiv.org/abs/1609.08042
[2] Khronos Group. 2017. OpenGL – The Industry's Foundation for High Performance Graphics. http://opencv.org/. (April 2017).
[3] NDN Project. 2017. Named Data Networking (NDN) – A Future Internet Architecture. https://named-data.net/. (April 2017).
[4] OpenCV team. 2017. OpenCV library. https://www.opengl.org/. (April 2017).
[5] Raspberry Pi Foundation. 2017. Raspberry Pi – Teach, Learn, and Make with Raspberry Pi. https://www.raspberrypi.org/. (April 2017).
[6] Ricoh. 2017. RICOH THETA. https://theta360.com/. (April 2017).
[7] Y. Sánchez, R. Skupin, and T. Schierl. 2015. Compressed Domain Video Processing for Tile based Panoramic Streaming using HEVC. In *Proceedings of IEEE International Conference on Image Processing (PV)*. 2244–2248.