

Frame by frame playback on the Internet video

Chikara Miyaji¹

Japan Institute of Sports Science,
3-15-1 Nishigaoka, kita-ku,
Tokyo, 115-0056, Japan

Abstract. Although frame-by-frame is one of the important operation to see sports movement precisely, it is not implemented in the streaming videos on the Internet. This article introduces Smart-method, a new method of frame-by-frame on the streaming videos. It is based on the combination of two methods, streaming and download; streaming is used for playback, the image download is used for frame-by-frame operation, and these operations are switched smoothly, users will not notice these switches. Smart-method is not only enables to see sports movement precisely it opens various applications possible on the Internet; attaching meta-data of the sports videos, flexible thumbnail, and editing (cut and join the videos) on the Internet.

Key words: streaming, frame-by-frame playback, and the Internet videos

1 Where the problem is?

It was almost 25 years has passed after computers have handled the videos, for example, Apple Inc. introduced QuickTime video technology in the year 1991. Currently no application exists to process videos without computers. In these 25 years, the Internet becomes wide spread and the video format has evolved from SD to HD, and even more to 4k/8k. The videos are shared on the Internet such as YouTube, and are browsed even on smart phones.

Does video on the Internet become easier to handle from these years? Adaptive bitrate streaming², the most recent technology for the Internet video, only handle play, pause and seek but no frame-by-frame playback. Because of the limitation, video editing still uses local video files. The Internet videos are not comfortable to observe sports movements precisely.

2 Why the Internet videos can't perform frame-by-frame?

First of all, it is not designed for these operations, but just for playback videos such as movie clips.

¹ chikara.miyaji@jpnnsports.go.jp

² In this article, the word streaming means Adaptive bitrate streaming.

On streaming, the data will not keep on the client after it played back, so it is difficult to perform step backward. Also if seeking will occur on streaming, the time to seek is not accurate because of the character of streaming. On streaming, the video file is divided into small fragments, and is downloaded sequentially to playback. Each fragment starts from the key-frame and followed by the difference frames. So, when seeking occurs, the client restrains to the key-frame instead of the real frame. This results that seek can't emulate frame-by-frame playback.

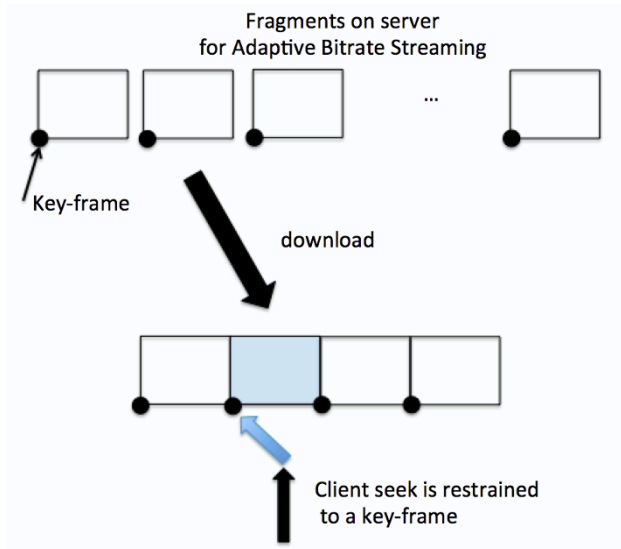


Fig.1: A client seek is restrained to a key frame on adaptive bitrate streaming.

The lack of frame-by-frame is not a lack of a function, but makes many applications impossible to perform on the Internet.

3 The purpose of the article

On this article, the author introduces new method of forward and backward frame-by-frame playback on the Internet, named Smart-method (Streaming Media Application enRichment Technology method), and explains its implementation. This method is general to any streaming method, so that it can apply on Apple's HLS or MS's smooth streaming. After frame-by-frame become possible on the Internet, new applications become possible. This article introduces new thumbnail methods based on Smart-method. Also a new video editing on the Internet is explained. It is amazing that Smart-method expands many possibilities of the Internet videos.

4 The implementation of frame-by-frame on Smart-method

Smart-method does not implement frame-by-frame on the streaming mechanism, instead, it implements frame-by-frame by downloading still image of the video form the server. If the still image is downloaded and overlapped on the video itself, it looks like the video player performs a frame-by-frame playback.

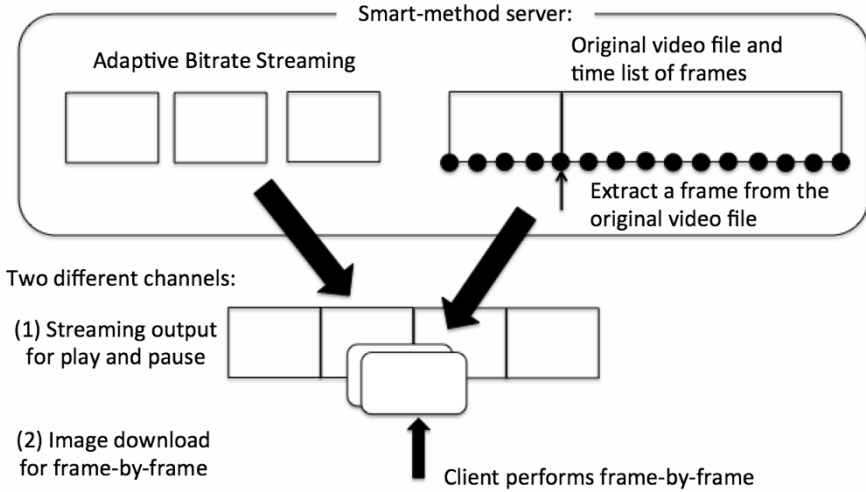


Fig.2: Smart-method uses two channels to play and to perform frame-by-frame.

Smart-method defines APIs to download the still image. There are two APIs for the server:

(1) API to get the list of the times (ms) of all frames of the video

`http://..server../dir/video/time_list.json => [0, 33, 67, 100, 133, ...]`

If the video is high frame-rate (ex. 120fps), the list should be [0, 8, 17, 26,...].

(2) API to get the image of the frame of a certain time (ms) of the video

`http://..server../picture.php?file=dir/video&ms=99 => an image`

The server program `picture.php`, a simple PHP application, returns an image. And inside the program, a `ffmpeg` command is running:

```
$ ffmpeg -ss 120 -i file.mp4 -f image2 -vframes 30 out%d.jpg
```

This command extracts 30 images of the frames of “file.mp4” start from 120 sec, which is enough to return the image.

If above APIs exist, frame-by-frame playback is done by the sequences below,

1. Pause the streaming video,
2. User choose forward step (or backward),
3. Get the current time of the video in ms,
4. Calculate the time of next frame from API (1),
5. Request the image of the time using API (2),
6. Display the image overlapped on the video

This is exactly what Smart-player (a video player in which Smart-method is implemented)[1] is doing. In the practical situation, it is difficult to set exact time only after pausing, thumbnail search (explain later) will be used before frame-by-frame playback.

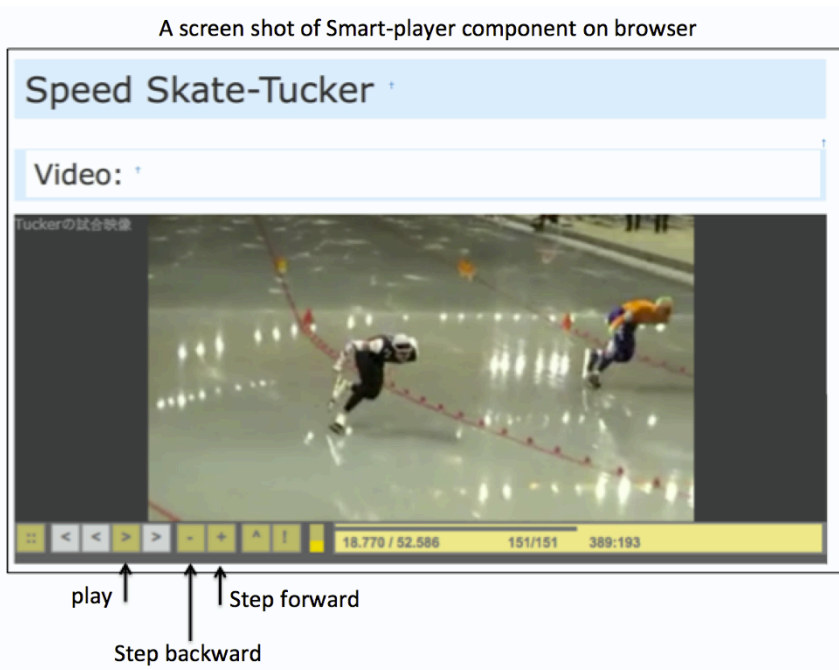


Fig.3: A screen shot of Smart-player component on browser.

The merit of Smart-method is free from the streaming protocols. For example, Smart-player for iOS (Using HLS) and Smart-player for PC (Using Smooth Streaming), both work exactly same on frame-by-frame playback.

The demerit of this method is the requirement of the programs on the server. The existing streaming service such as YouTube can't perform Smart-method.

5 Applications of Smart-method

When frame-by-frame become possible, several applications expand its functionalities. Some of the applications explain here.

5.1 For Sports video browsing

The original intention to expand video player came from the poor playback of YouTube player. For browsing sports movement, it is very important to see the exact point such as ball release or hitting. Also it is important to use frame-by-frame to understand very fast movement. Although many applications can perform frame-by-frame on local video files, no application performs frame-by-frame on the Internet video. Smart-method helps the user to see the Internet video precisely.

Meta-data is the base of searching the scenes, or analyzing the game. For example, Major League Baseball Advanced Media, one of the top company of sports media industry in US, uses more than 500 peoples for attaching meta-data on their baseball videos. It means that human eyes are important even on the most advanced company. There are two important requirements for attaching meta-data on sports video,

- Meta-data requires exact timing of the event
- Human power is necessary to define the meta-data

Smart-method helps these situations; frame-by-frame helps to find exact time of the event and the use of the Internet helps to share the work with many peoples.

5.2 Implementation of thumbnail

For sports video browsing, it is common to see some specific scene instead of watching it from the beginning to the end. Thumbnails, 100 to 200 selected images from the video, are used to find the important scene quickly. For this application, API to download a image is used, and lower image resolution option is added on this API to make download faster. Although Smart-player takes about 7.5 sec to download 150 thumbnails, it will not stop the user operations because it is running as background.

The list of the thumbnail does not need to be laddered, but the time intervals can be varied from sparse to dense. Usually videos of sports will have unimportant scenes and important scenes according to the content, and the thumbnail should reflect these characteristics; unimportant scenes will have sparse intervals, and the important scenes will have dense. This mechanism is called “non-linear thumbnail”.

Non-linear thumbnail helps to find the important scenes quickly, but sometimes, each user will have different importance in sports, for example coach’s view and judge’s view are different. For such case, Smart-player provides “on-demand thumbnail” to download dense thumbnails according to the period where user select.

If thumbnails of all frames are collected in some period, the action of dragging the mouse (for PC) or moving the finger on the screen (for smart phone) makes very smooth movement of thumbnails. Action moving left to right make normal play, and the action moving right to left make backward play, and repeating both actions help to

see the movement again and again. This operation is very useful to examine the movement.

The flexibility of the API of Smart-method makes various thumbnail methods possible.

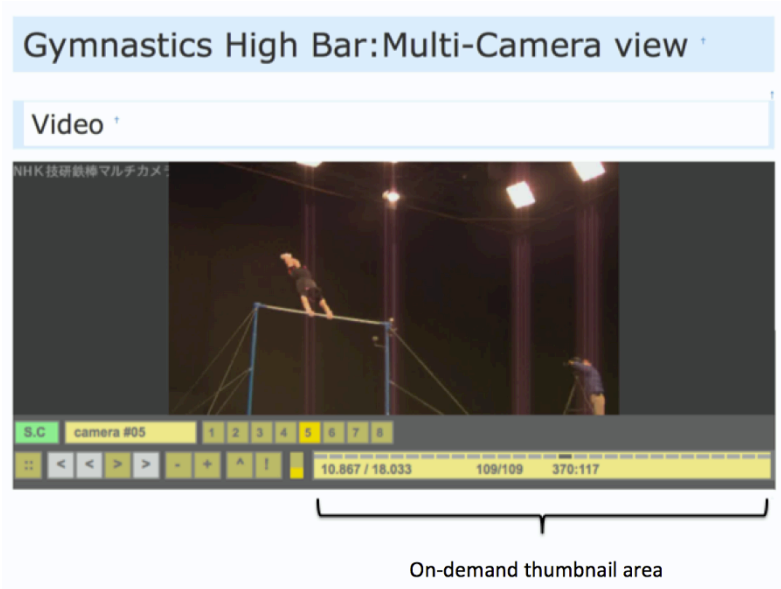


Fig.4: A snap shot of using on-demand thumbnail

5.3 Implementation for Editing

Streaming uses metafiles to define fragments structure. Based on these metafiles, cut and join are realized by editing of these metafiles. Because of the time period of the fragments, the smallest unit of editing is restricted on its size. For this restriction, frame-by-frame editing based on metadata looks impossible.

Fig.5 shows simple schematic representation of join operation of video A and video B. If user wants to join “part of video A” and “part of video B”, the joined metadata will include unnecessary parts.

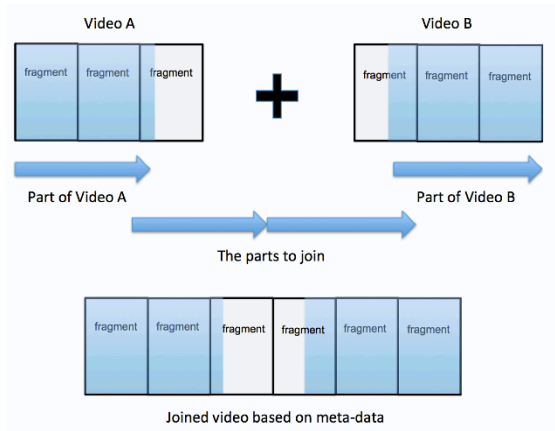


Fig.5: A schematic representation of join operation based on meta-data.

On Fig.6, frame data are shown and blacks are the part to join. For the frame data of video A+B, a joined lists of both black part of video A and video B will be created.

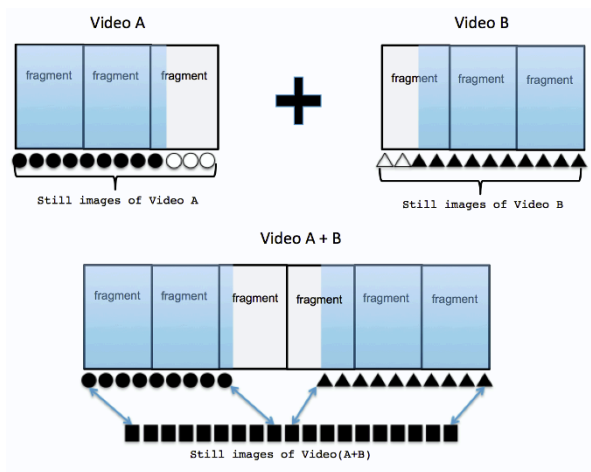


Fig.6: A schematic representation of joined video with frame data.

If user will check the video A+B by frame-by-frame playback, the frame data will show exactly what user edited and will not include unnecessary part. If the user will check the video A+B by replay, unnecessary part will be replayed but that part will be too short to notice.

Here are my assumptions:

- When playing the video, it will not important to include small unnecessary part
- When doing frame-by-frame playback, it is important to see accurate frame editing

Smart-method satisfies these assumptions, and it means that video editing over the Internet is not impossible.

This kind of editing will be useful for simple sports videos editing, such as collect batting scenes or cut some scenes. It will be also useful to use it as proxy editing for professional use, because it includes precise frame-by-frame data.

When 4K/8K videos will become popular, all videos will be uploaded, and editing this way will be realistic.

5.4 The possibility for Sports analysis

Almost all current sports video analysis applications will use local video file as the input, because these applications require frame images for their image processing. Smart-method opens the way to get frame image via the Internet. If these applications will include Smart-method mechanism for their input, these applications will have more choices for their input. Also if these applications will be put on the Internet, these applications will become services. This will make more choice for the users.

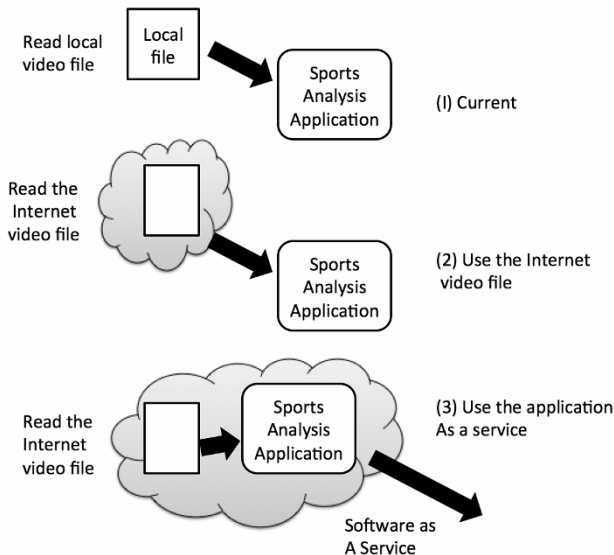


Fig. 7. Three steps to evolve Sports Applications

References

[1] Chikara MIYAJI: A New Browsing Method for Sports Video on the Internet, Proc. of the 8th International Symposium on Computer Science in Sport (IACSS2011), pp54-57, 2011