June 2020

Interactive Videos with Embedded High-Resolution Images

Anonymous

Interactive Videos with Embedded High-Resolution Images

ABSTRACT

Currently, videos generally cannot be zoomed in or panned across. Videos capture motion but not detail; still images do not capture motion, but can be very detailed. The techniques described herein unite the motion-capturing features of videos with the detail-capturing features of still images to enable a user to zoom in, pan across, listen to narrations, and generally interact with frames in a video. In this manner, videos are provided that are interactive and engaging.

KEYWORDS

- Interactive video
- Zoomable video
- Panning
- User engagement

BACKGROUND

Currently, videos generally cannot be zoomed in or panned across. Videos capture motion but not detail; still images do not capture motion, but can be very detailed. A combination of substantial level of detail, as offered by high resolution still images, and video that includes motion can provide a richer experience by enabling pan/zoom operations in the high resolution image while also providing motion.

DESCRIPTION

This disclosure describes techniques that enable a user to zoom in, pan across, listen to narrations, and generally interact with frames in a video. Traditionally, videos capture motion but

not detail; still images do not capture motion, but can be very detailed. The described techniques unite the motion-capturing features of videos with the detail-capturing features of still images.

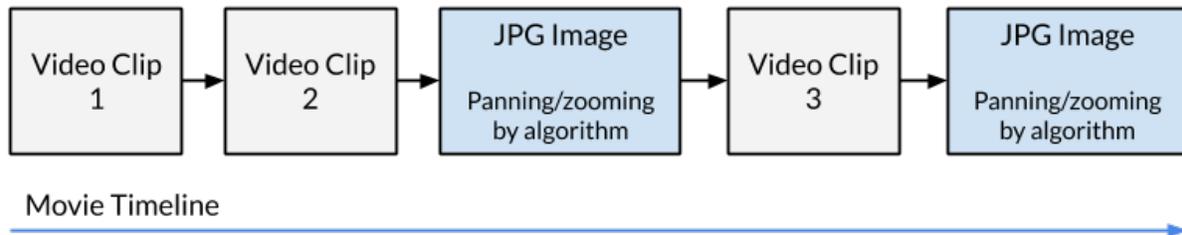*Creation of an interactive, e.g., zoomable or pannable, video*

Fig. 1: Timeline of compiled movie that includes video clips and still images

Per the techniques, illustrated in Fig. 1, a video compilation (movie) includes video clips, e.g., clips 1, 2, and 3, that are seamlessly blended in with high-resolution still images. The resulting video is stored in a blended format, e.g., incorporating elements from video standards such as MPEG and still-image standards such as JPEG. The video includes audio layers both for the video clips and the still image.

A pop-up icon or other user-interface element is provided that identifies the video frames (or portions) that have corresponding high resolution still images. Users can freeze the video at one or more of the interspersed images, zoom into the image, pan across the image, listen to narrations that accompany the image, etc. The user can also use a seek bar to jump to different portions of the video.

An interactive video can be created in any of the following ways:

- **Manual packaging**, where the video-creator plays a role similar to a film editor or director, splicing together video clips and still images based on artistic considerations. Although in most cases it is expected that the interspersed still images are semantically

correlated with the video clips, this need not be true for all content. The decision to relate still images to the video clips is an artistic one that is left to the video creator.

- ***Automatic collation***, which uses a camera that is capable of simultaneously shooting movies and high-resolution still images.

The accompanying audio is stored as a separate channel of the blended format. Multiple audio channels, e.g., one carrying music, another carrying commentary, etc. can be included. The audio channels can be recorded at the same time as the video or added during video editing. When a viewer freezes the video to examine a frame in greater detail by zooming or panning, audio narration relevant to the frozen frame can be started.

Raw JPEG images are typically 2-8 megapixels in size while raw MPEG videos can have bitrates of 4-6 Mbps (at a resolution of 1080×720 pixels at 60 frames per second). These rather large sizes are not typically uploaded to video sharing or social media services. Rather, a substantially compressed version of such images or videos is uploaded. Per the described techniques, a video creator can select sections of the video (up to the entire video) for which original, high-resolution, video frames and still images are uploaded. The high-resolution video or imagery enables zoom and pan functions during playback of the video.

For example, details that become visible at high resolution can be live-streamed during zoom-in. The interspersing of movie frames and still images is seamless to the user such that transitions between viewing a frozen video frame or an interspersed still image are not perceivable. If the user uploads the generated content to a website or application that does not support the interactive video format described herein, the uploaded video plays like a regular, e.g., non-zoomable, non-pannable, video.

Playback of an interactive, e.g., zoomable or pannable, video

During playback of the video, a displayed icon serves to inform the viewer of a zoomable or pannable section. Alternatively, an icon can be used to inform the viewer of sections that are *not* zoomable or pannable. Viewers can freeze the video at one or more of the interspersed images, zoom into the image, pan across the image, listen to narrations that accompany the image, etc. Viewers can also use a seek bar to jump to different portions of the video. The seek bar can also include icons that indicate the available high resolution images within the video.

For example, while the video is being played on a touchscreen device, the video can automatically pause when the viewer pinches-to-zoom and automatically restart when the viewer zooms out all the way. When a viewer freezes the video to examine a frame in greater detail by zooming or panning, audio narration or music relevant to the frozen frame can be started. At any time, the viewer can jump off the interactive-video experience, e.g., restart playback of the video.

The blended format video as described herein can be viewed on various devices, e.g., desktop, mobile device, tablet, television, virtual reality headset, etc. The blended format can be used by video sharing, social media, blogging, and/or photo sharing services, etc.

The blended format described herein enables the creation of dynamic videos, e.g., videos with sections that are pre-recorded and sections that are live-streamed. An example of a dynamic video is that of a newscaster reporting stock market news. The section of the video that includes the newscaster may be a few hours old, while the section that shows the stock market graph can be livestreamed such that it is always current, e.g., real-time. Another example of dynamic video is one where an event, e.g., a concert performance, is filmed from multiple angles, and the corresponding content is included in the blended format. A viewer can then change viewing angle at will or zoom into interesting features from the one or more video feeds.

In this manner, sections of a video are made amenable to zooming and panning, and, depending on the resolution of the interspersed images, to inclusion of substantial detail. The resulting videos, provided in a blended format, are more interactive and engaging than traditional videos.

CONCLUSION

Currently, videos generally cannot be zoomed in or panned across. Videos capture motion but not detail; still images do not capture motion, but can be very detailed. The techniques described herein unite the motion-capturing features of videos with the detail-capturing features of still images to enable a user to zoom in, pan across, listen to narrations, and generally interact with frames in a video. In this manner, videos are provided that are interactive and engaging.

REFERENCES

[1] Roth, James M., and Kenneth G. Oetzel. "End-user-navigable set of zoomed-in images derived from a high-resolution master image." U.S. Patent Application 12/045,629, filed Mar. 10, 2008.

[2] Trimeche, Mejdi, and Sami Vaskuu. "System and method for implementing improved zoom control in video playback." U.S. Patent Application 11/615,597, filed Dec. 22, 2006.

[3] Kim, Sin-ae, Hyun-mi Park, Deok-Hee Jeong, and Joo-Kyung Woo. "Image display apparatus and method of controlling the same via progress bars." U.S. Patent 8,762,844, issued June 24, 2014.

[4] Nickolov, Radoslav Petrov, Lutz Gerhard, Ming Liu, Raman Narayanan, and Drew Edward Steedly. "Multi-layer image composition with intermediate blending resolutions." U.S. Patent 8,619,083, issued December 31, 2013.

[5] Casagrande, Steven Michael. "Methods and apparatus for displaying information regarding interstitials of a video stream." U.S. Patent 8,165,451, issued April 24, 2012.