June 2022

Upscaling Image Resolution Using Optical Zoom and Rotation

Dmitrii Boiarshinov

Upscaling Image Resolution Using Optical Zoom and Rotation

ABSTRACT

This disclosure describes techniques to capture images or videos of high resolution with cameras and optical components of ordinary size and specification as found in typical webcams or smartphones. Per the techniques, one or more cameras perform an optical zoom to sections of the field of view, capturing high-resolution sub-images of the field of view. The cameras raster across the field of view, rotating as necessary, to capture high-resolution sub-images that are integrated to reconstruct a high-resolution image of the entire field of view. As the scene changes, the cameras refocus and re-capture the sections of the field of view that have experienced change, such that the image as a whole remains updated.

KEYWORDS

- Video conferencing
- Digital photography
- Image resolution
- Optical zoom
- Image reconstruction
- Image merging
- Image stitching
- Photo montage
- Image rasterization
- Streaming video

BACKGROUND

There is longstanding interest in achieving high resolution in digital photography and videography, as higher resolution provides better image/video quality and greater detail. Achieving high resolution has thus far entailed the use of high pixel density in the CCD/CMOS / photon sensor; large lenses to gather greater amounts of light; and/or high-quality optical components. These in turn reduce portability, lead to large image file sizes, require more intensive computation, and are expensive.

DESCRIPTION

This disclosure describes techniques to capture images or videos of high resolution with cameras and optical components of ordinary size and specification, e.g., as found in typical webcams or smartphones. Per the techniques, one or more cameras optically zoom into sections of the field of view, capturing high-resolution sub-images of the field of view. The cameras raster across the field of view, rotating as necessary, to capture a sequence of high-resolution sub-images. The sub-images are integrated to reconstruct a high-resolution image of the entire field of view. As the scene changes, the cameras refocus and recapture sections of the field of view that have experienced change, such that the image as a whole remains up to date. The cameras can be of quality normally found in smartphones or webcams, so long as they can rotate slightly and optically zoom while staying in place.

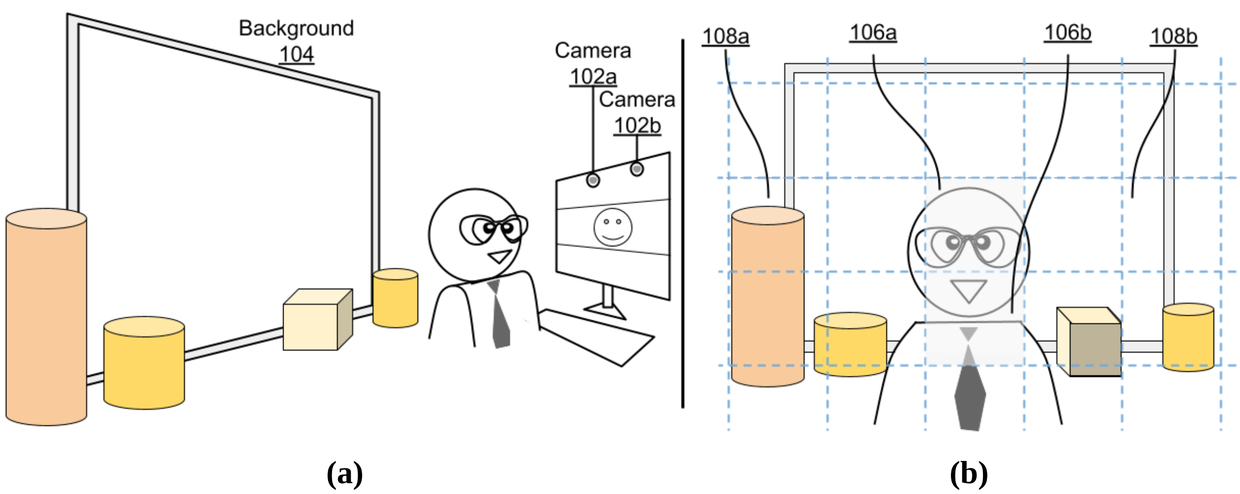

Fig. 1: (a) A user on a video conference; (b) View from the user's camera(s)

Fig. 1 illustrates an example of upscaling the resolution of images or videos using optical zoom and rotation. Fig. 1(a) illustrates a user participating in a video conference, an example application of the techniques described herein. The user's device has two cameras (102a-b) and is set against a background (104). The view from the user's camera(s) appears in Fig. 1(b).

The camera(s) take an initial image with no zoom. The field of view is sectioned. For example, as illustrated in Fig. 1(b), the sections can be rectangular. Each section is optically zoomed into and captured at high resolution. If necessary, the camera can be rotated to focus on particular sections, e.g., edge sections. Zoomed-in shots of sections are merged together to form a high-resolution, composite image or video. The merging (also known as stitching) of sections can be performed by matching features with the initial shot. Further, the rotation angle and zoom can inform the coordinates at which the zoomed-in section is to be embedded within the initial image.

Upscaling an image or frames of a video is done dynamically. Sections of the image with greater detail or faster rates of change can be dwelled upon for greater amounts of time or updated at greater frequency. In the example of Fig. 1(b), which is a talking-head type scenario, the cameras can spend relatively less amounts of time updating the background sections (108a-b) and greater amounts of time updating sections that include the user's head (106a-b). Indeed, for talking-head scenarios, it can be sufficient to zoom in and update the sections of the image including the speaker and embed changes to those sections in the initially captured static image.

Similarly, for video conferences, which typically take place in rooms with static backgrounds, the initially captured image can be upscaled and sent once to the receiver, with updates being sent for sections that depict conference participants. Reconstruction of the entire field of view can be done at the receiver. Doing so reduces bandwidth while maintaining high resolution. To maintain a fast update rate, changes in the background can be updated immediately at low resolution, with a higher-resolution scan done shortly thereafter.

When multiple cameras are available, one of the cameras can statically maintain view on the entire visual field while other cameras zoom into particular sections, rotating as necessary to better target regions of focus. For example, in a room full of people, changes can be captured under normal resolution by a static camera and an upscaled resolution image of the currently speaking person can be captured by zooming into and/or rotating another camera towards the person. In this manner, unchanging areas have high resolution, dynamically changing areas have normal resolution, and the currently speaking person has upscaled (high resolution). Alternatively, a single camera can multitask between capturing the entire field of view and capturing zoomed-in sections.

A camera can optically zoom out periodically to detect regions that have experienced change. Regions with change can be zoomed in digitally or optically such that the entire field of view remains updated. Optical zoom can enable different sections of the zoomed image to have the same upscaled resolution since the sections share the camera matrix.

The dimensions of sections can be adjusted to match the amount of detail, e.g., small sections can be used for regions of the image with greater detail and vice versa. An object, face, or motion detector can be used to determine sections of the image to be updated. Sections can be zoomed in using a raster-like schedule (left-to-right and top-to-bottom) or from areas of greater detail to areas of lesser detail. To further improve image or video quality, small changes in the angle of the shot can be used to recreate 3D scenes.

The described techniques apply generally to the capture of images and video, and particularly to the capture of videos with a static background and actively changing foreground, e.g., talking-head style videos as often found in video conferencing or video chat applications.

CONCLUSION

This disclosure describes techniques to capture images or videos of high resolution with cameras and optical components of ordinary size and specification as found in typical webcams or smartphones. Per the techniques, one or more cameras perform an optical zoom to sections of the field of view, capturing high-resolution sub-images of the field of view. The cameras raster across the field of view, rotating as necessary, to capture high-resolution sub-images that are integrated to reconstruct a high-resolution image of the entire field of view. As the scene changes, the cameras refocus and re-capture the sections of the field of view that have experienced change, such that the image as a whole remains updated.