December 28, 2018

Automatic Removal Of Artifacts In 360° Videos

Anonymous

Automatic Removal Of Artifacts In 360° Videos

ABSTRACT

Many 360°videos captured using a camera rig and tripod setup have artifacts related to the tripod that the rig is attached to. This disclosure describes automatic removal of such artifacts by use of machine learning techniques. Portions of the video that include artifacts are detected and removed automatically using trained machine learning model(s). The ML models are trained to recognize production equipment based on videos captured in neutral environments or based on markers on the production equipment. The detected artifacts are automatically replaced based on the context and ensuring that the resultant video is stable.

KEYWORDS

360° video; 360° photo; 360° camera; 360 degree; camera rig; tripod; automatic correction; artifact removal; image reconstruction; video production; virtual reality; machine learning

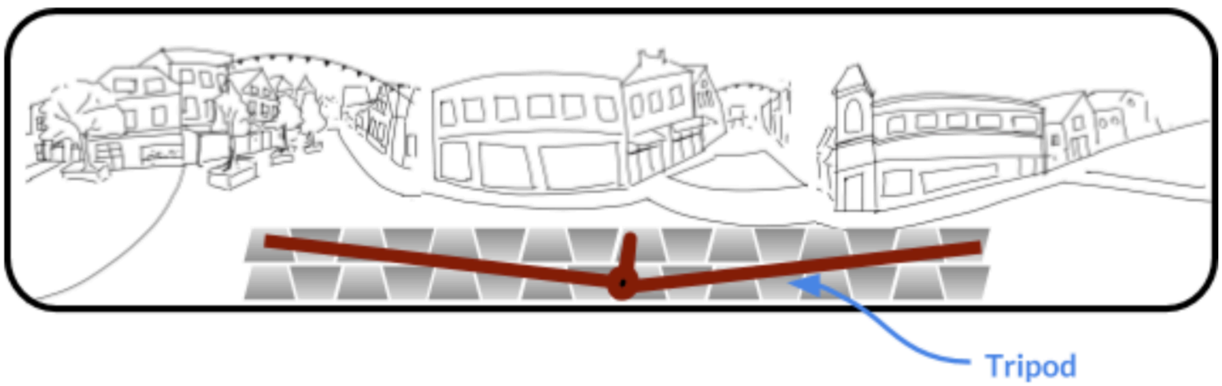
BACKGROUND

Capture of 360°videos is often performed using a camera rig that can record video simultaneously in 360°. Camera rigs are typically mounted on a tripod. Many 360°videos captured using a camera rig and tripod setup have artifacts related to the tripod that the rig is attached to.

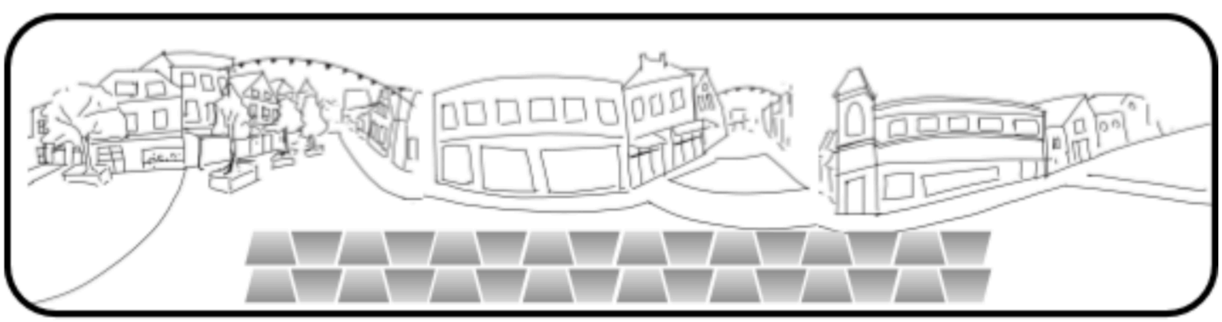
When viewing such videos, e.g., via wearable devices such as goggles, looking down can often lead to a break in the immersive experience. Video producers take substantial effort to eliminate camera rigs from the captured video, e.g., by blanking out that portion of the footage, by capturing additional footage that is edited into the final result, etc. Video content hosts, e.g., video-hosting websites, benefit from higher quality 360°videos that do not include artifacts.

DESCRIPTION

This disclosure utilizes machine learning based image correction techniques to automatically remove artifacts related to the camera rig or other production equipment. Portions of 360°video footage that include the production equipment are automatically filled in with content that fits the context and that doesn't distract from the directed experience.



Tripod shows up as an artifact in captured 360° image



360° image after automatic artifact removal

The top portion of the above figure illustrates an example of the use of such techniques. As can be seen in top portion of the figure, a spherical image (e.g., a frame of a 360°video) captured using a camera mounted on a tripod includes an artifact in the form of the tripod that is set on top of concrete blocks. Note that the above image is shown in rectangular shape for ease of illustration; the actual shape of a 360° image may be different, e.g., spherical.

Machine learning techniques are utilized to replace the tripod at the bottom of the spherical image with concrete blocks (or portions) that fits in the context, as illustrated in the bottom portion of the figure.

An example process for such replacement is as follows:

1. Capture 360°video.
2. Obtain information regarding the footage (e.g., quantity and location of the footage in a frame) directly below the camera that is to be replaced, e.g., as an angle.
3. Apply machine learning techniques to fill the identified portion of the footage, e.g., fill in the portion below the camera with generated content. The ML techniques then reconstruct image data for the portion of the image below the defined angle, e.g., based on other parts of the image.

The process described above can be further enhanced. For example, the camera rig that is utilized to capture the video can be modified to include markers or colors that allow easy identification of the portions to be replaced. The machine learning model(s) used to replace the identified portions can be trained with images of the different types of production equipment utilized for capture that are obtained in neutral environments. This allows the ML model to learn the types of artifacts that need replacement. The ML techniques can also be trained to automatically identify and remove production equipment from the captured footage.

Generated image data, e.g., multiple frames of video, is further processed to ensure that it is stable over time so as to not distract from the video experience. Such processing can be performed after artifact removal, or multiple frames can be processed simultaneously for artifact removal.

Such automated artifact removal can be employed in tools for production of 360° video, and for automatic processing of videos submitted online by users, e.g., to a video-hosting website, social media website, etc. The technology can also be used for static images, e.g., as an automatic tripod removal tool in 360°photos.

Image reconstruction - removal of undesired objects and replacement with contextual fill - has been demonstrated already on static images. This disclosure extends image reconstruction techniques, e.g., as described in [1], to work automatically as part of a 360°video production workflow. Automated detection and replacement in this manner can help lower production costs, e.g., the cost to shoot extra footage, post-production costs to remove artifacts, etc.

CONCLUSION

Many 360°videos captured using a camera rig and tripod setup have artifacts related to the tripod that the rig is attached to. This disclosure describes automatic removal of such artifacts by use of machine learning techniques. Portions of the video that include artifacts are detected and removed automatically using trained machine learning model(s). The ML models are trained to recognize production equipment based on videos captured in neutral environments or based on markers on the production equipment. The detected artifacts are automatically replaced based on the context and ensuring that the resultant video is stable.

REFERENCES

1. https://www.sciencealert.com/nvidia-neural-network-editing-program-for-irregular-inpainting accessed Dec 24, 2018