December 12, 2017

Classifying Aspect Ratios Of Images

Andrew Gallagher

CLASSIFYING ASPECT RATIOS OF IMAGES

ABSTRACT

Disclosed herein is a mechanism for detecting and/or adjusting aspect ratio in images and videos. For example, the mechanism can include analyzing a digital image having pixels by identifying sets of pixels corresponding to one or more objects in the digital image, analyzing the set of pixels corresponding to at least one object in the image to determine an estimated image aspect ratio, and applying an image transform to the digital image based on the estimated image aspect ratio to generate a modified digital image. In a more particular example, the one or more objects can correspond to faces, animals, logos, signs, or vehicles. In another example, the mechanisms can include generating a quality rating for the digital image based on the estimated image aspect ratio.

BACKGROUND

Images and videos are widely available online. For example, images are often returned in response to search queries, and videos are available for streaming via different video content providers. Groups of images and videos are often presented to users, for example, in response to a search query, and the images and videos are typically ranked prior to presentation. However, the images and videos may be distorted, for example, an image or a video may have an incorrect aspect ratio. For example, a 4:3 aspect ratio video can be used to fill a 16:9 display and subsequently re-encoded. It can be difficult to consider incorrect aspect ratios when selecting images and videos for presentation to a user. Additionally, presenting a distorted image or video can be unpleasant to a viewer, thereby reducing the viewing experience.

DESCRIPTION

An image provider or a video content provider can use the mechanism to analyze an image or a video to determine if the image or the video is distorted. In particular, the mechanism can be used to analyze the image or the video to determine if the image or the video has been encoded with an incorrect aspect ratio, thereby causing the image or the video to appear to be narrow or to appear to be wide. The mechanism can then score the image or the video based on whether the aspect ratio is incorrect, and the mechanism can use the score when ranking images or videos, for example, in determining which images or videos to return in response to receiving a search query. Additionally or alternatively, the mechanism can, based on an estimated aspect ratio, transform or otherwise modify the image or the video to have a corrected aspect ratio. The aspect ratio can be determined using a machine learning algorithm that classifies whether an object (e.g., a face, a person, an animal, a logo, a sign, a vehicle) detected in the image or in a frame of the video is narrow or wide relative to a corresponding template of the object.

FIG. 1 illustrates an example method for classifying images based on aspect ratio. The method can be performed by a system that stores or provides images and/or video content, such as a server that stores images and serves images in response to receiving a search query, a server that hosts and delivers video content, and/or any other suitable type of system.

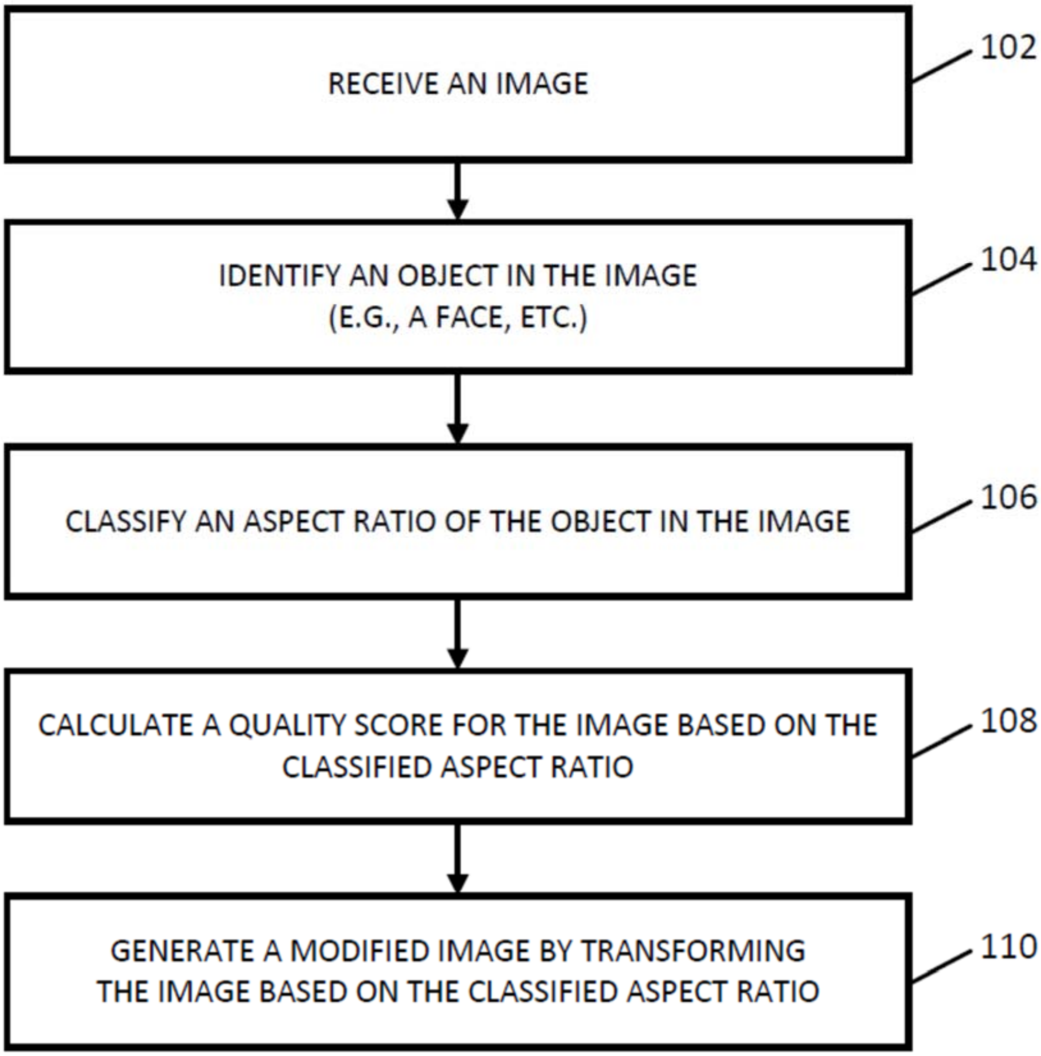


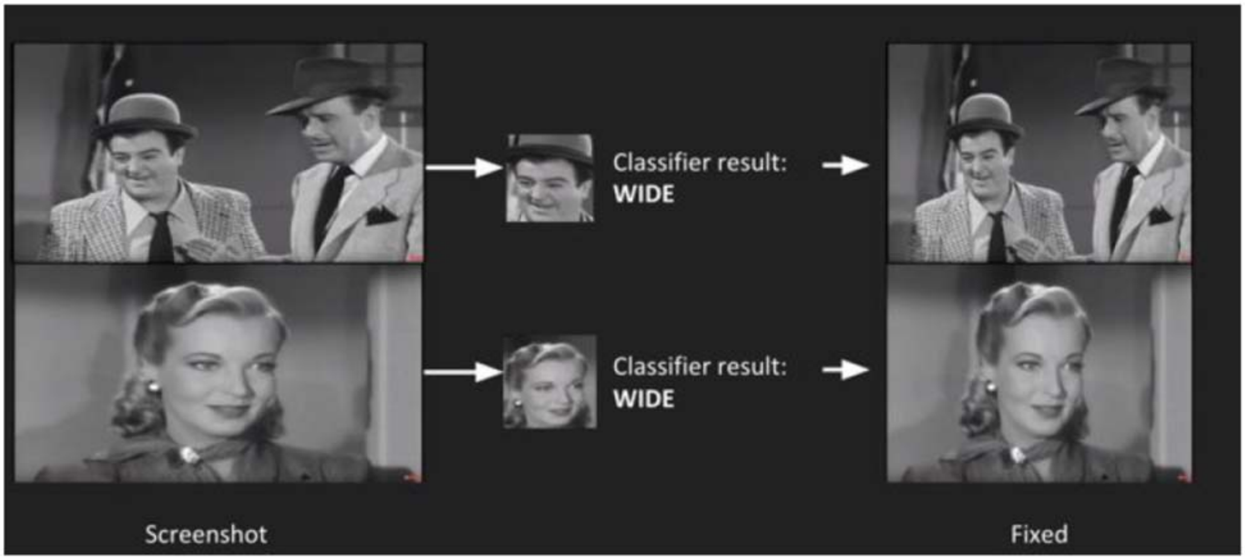
FIG. 1

At step 102, the server can receive an image to be classified. In some cases, the image can be a frame of a video stored on the server. For example, in some cases, the video can be a video uploaded to the server by a user for storage and delivery to other users.

At step 104, the server can identify an object in the image. The object can be any suitable object, such as a face of a person, a body of a person, an animal, a stop sign, a fire hydrant, a vehicle, and/or any other suitable type of object that has a typical shape or a semi-rigid shape (e.g., a typical height-to-width ratio, and/or any other suitable typical shape). The server can identify the object in any suitable manner, for example, using any suitable combination of image recognition techniques. As a more particular example, the server can use an object detector that determines a type of object in the image or whether a particular type of object is present in the image and a portion of the image that contains the object. As another more particular example, the server can match a template of a face to an object in the image. As yet another more particular example, the server can use edge detection to identify an outline of an object in the image. In instances where the server identifies a portion of the image that contains the image, the portion of the image can be indicated with, for example, a bounding box or a mask of pixels that indicate a region of the image that contain the object.

At step 106, the server can classify an aspect ratio of the image. The aspect ratio can be classified using a machine learning algorithm that has been trained on training samples of the same type as the detected object. For example, in instances where the detected object is a face, the algorithm can be trained on any suitable number (e.g., a thousand, two thousand, and/or any other suitable number) of example images of faces. Each training sample object can be paired with a classification that indicates the aspect ratio of the training sample object, which the server can use to classify the aspect ratio of the image. Any suitable type of classifier can be used, such as a deep convolution model.

For example, as shown in the screenshot below, the server can extract a face (or any other suitable object) in image frames from a video and analyze the face using the classifier to determine that the face appears to be "WIDE" with a stretch factor of 4/3 compared to unstretched faces.



In another example, as shown in the screenshot below, the server can extract a face (or any other suitable object) in image frames from a video and analyze the face using the classifier to determine that the face appears to be "NARROW" with a stretch factor of 3/4 compared to unstretched faces.

The output classification can be in any suitable format. For example, the output classification can be a stretch factor of the detected object. As a more particular example, in instances where the classifier determines that the detected object is too narrow (e.g., has 3/4 the width of an unstretched object of the same type), the classifier can output a correction factor indicating a scaling required to correct the width of the detected object. As a specific example, in an instance where the classifier determines that the detected object has 3/4 the width of an unstretched object of the same type (e.g., in an instance where an image with a 16:9 aspect ratio has been squeezed for presentation in a 4:3 display), the correction factor can indicate that the image should be horizontally scaled by 4/3 to correct the width of the detected object. Note that, in some instances, the classifier can additionally output a confidence of the classification. Additionally or alternatively, in some instances, the classifier can classify the detected object multiple times and can select a single estimate for the aspect ratio by, for example, selecting the aspect ratio with the highest estimated confidence, the average aspect ratio, the median aspect ratio, and/or in any other suitable manner.

Note that, in some instances, the aspect ratio classification can be performed without first detecting an object. For example, in some instances, the image can be used as an input to a machine learning algorithm that has been trained to estimate an aspect ratio of the image, where the training samples are not detected objects, but an entire image.

At step 108, the server can calculate a quality score for the image based on the classified aspect ratio. The quality score can be calculated using any suitable information. For example, the quality score can be based on the correction factor determined by the classification algorithm at step 106. As a more particular example, the quality score can be calculated such that correction factors most different from 1 (that is, requiring more correction) receive a lower quality score. Note that, in instances where the image is a frame of a video, a quality score can be calculated for multiple frames of the video by aggregating individual scores for different frames of the video (e.g., as a mean quality score, a weighted average of quality scores, and/or in any other suitable manner). An aggregate quality score can encompass any suitable portion of the video, such as an introduction section corresponding to a first *N*% of the video, or can describe the quality of the entire video. In some instances, the quality score can be used for any suitable purpose. For example, the server can use the quality score to rank images or videos prior to presenting them to a user.

At step 110, the server can generate a modified image by transforming the image based on the classified aspect ratio. An example of a distorted image 200 is shown in FIG. 2A. As shown, the object (a face) appears too wide.

FIG. 2A

The server can generate a modified image 250 as shown in FIG. 2B by calculating a correction factor as described above in connection with block 106, and can apply a horizontal scaling with the correction factor to distorted image 200 to generate modified image 250. The server can then save modified image 250 for future use (e.g., to present to a user in response to receiving a search query that retrieves the image as a result, and/or for any other purpose). In instances where the image corresponds to a frame of a video, the server can save a modified frame, and can then proceed to analyze a next frame of the video in a similar manner as described above.

FIG. 2B

Accordingly, a mechanism for detecting and/or adjusting aspect ratio in images and videos is provided.