

## Technical Disclosure Commons

---

Defensive Publications Series

---

April 18, 2016

# IMAGE ORIENTATION DETECTION

Krzysztof Kulewski

Sevket Babacan

Follow this and additional works at: [http://www.tdcommons.org/dpubs\\_series](http://www.tdcommons.org/dpubs_series)

---

### Recommended Citation

Kulewski, Krzysztof and Babacan, Sevket, "IMAGE ORIENTATION DETECTION", Technical Disclosure Commons, (April 18, 2016)

[http://www.tdcommons.org/dpubs\\_series/190](http://www.tdcommons.org/dpubs_series/190)



This work is licensed under a [Creative Commons Attribution 4.0 License](https://creativecommons.org/licenses/by/4.0/).

This Article is brought to you for free and open access by Technical Disclosure Commons. It has been accepted for inclusion in Defensive Publications Series by an authorized administrator of Technical Disclosure Commons.

## IMAGE ORIENTATION DETECTION

### ABSTRACT

An orientation detection system sets an orientation of an image or a video based on the detection of an orientation of objects present within the image or video. The system detects the presence of one or more objects within the image or a frame of the video. The system determines spatial orientation parameters for the detected objects in the image. The system then constructs a histogram with a plurality of bins based on the determined spatial orientation parameters. The system counts the number of objects in each of the bins. The system determines the highest number of objects ( $M$ ) in a single bin from the plurality of bins. Thereafter, the system determines whether there is only one bin with a number of objects greater than  $M/2$ . The system then determines the orientation of the image based on spatial orientation parameters for the bin with the number of objects greater than  $M/2$ . If there are multiple bins with a count of objects greater than  $M/2$ , the system does not change the orientation of the image.

### PROBLEM STATEMENT

People build large collections of personal photographs taken over the years. A vast majority of digital cameras available in the market have image sensors designed for taking photos in landscape orientation. At the same time, landscape orientation is not always the best choice from a scene composition point of view, and a user often rotates the camera to obtain portrait orientation. Some digital cameras have a built-in orientation sensor that is able to detect device orientation and use that information while saving the image. However, due to the fact that

not all cameras have such sensors, e.g. older and inexpensive, but widely popular, cameras, this results in a huge bundle of photos which were and are taken in portrait mode but then saved as landscape images. The user may then have to manually rotate each image taken in portrait orientation to achieve the desired display effect, which is particularly time consuming if the user is working with or uploading a large batch of images from their old collections. There are opportunities to automatically detect the orientation for an image captured with an image capturing device lacking orientation sensor data.

### DETAILED DESCRIPTION

The systems and techniques described in this disclosure relate to an orientation detection system to detect and correct the display orientation of visual digital media. The system can be implemented locally on a client device or implemented across a client device and server environment. The client device can be any electronic device such as a mobile device, a smartphone, a tablet, a handheld electronic device, a wearable device, an image capturing device, a camera, a camcorder, video camera, etc.

Fig.1 illustrates an example method 100 to detect orientation of an image or a video. The method 100 can be performed by an orientation detection system. The system detects 102 one or more objects in an image or a frame of a video. The one or more objects in the image may include common known types of objects, e.g., human faces, text, animals, cars, etc.

The system determines 104 spatial orientation parameters of the detected objects in the image. The spatial orientation can be described in terms of three angles, namely, a roll angle, a Pan (Yaw) angle, and a tilt angle, each with range of -180 to 180 degrees. The roll angle

indicates how much clockwise/anticlockwise the object is rotated relative to an image vertical and about the axis perpendicular to the object. The Pan (Yaw) angle indicates how much leftward/rightward the object is pointing relative to the vertical plane perpendicular to the image. Further, the tilt (pitch) angle indicates the extent upto which the object is pointing upwards/downwards relative to the image's horizontal plane. A "0" orientation is considered to be the natural orientation for the object, e.g, a face looking straight into the camera, text in usual written orientation, car from front, etc. Further, for all the types of objects detected in the image, the system checks whether they are orientated in a similar way in all the situations, e.g., for an image containing text, the text is usually neither upside down nor vertically rotated.

In another embodiment, the system may determine the spatial orientation parameters of the objects in a video. The system determines the orientation of objects in each frame or each  $n^{\text{th}}$  frame of the video. The system analyzes each  $n^{\text{th}}$  frame to detect different orientations for consecutive probing points. The probing points may include certain reference points on each frame designated by the system in order to detect the orientation. If a different orientation is detected for different probing points, an additional probing can be used to determine an exact frame where the orientation changes. Further, the system determines the roll angle for each of the plurality of objects in the frames and assigns detection probability scores to each object and ignores any non-likely matches from further processing. Thereafter, each fragment of the video can be rotated separately in an appropriate manner.

The system constructs a histogram with a plurality of bins based on the determined spatial orientation parameters. Specifically, the system collects the roll angles of the detected objects to construct the histogram. In an example, the histogram has 4 bins corresponding to

ranges of roll angles:  $[-45, 45)$ ,  $[45, 135)$ ,  $[-135, -45)$ ,  $[135, 180) \cup [-180, -135)$ . The image histogram may be produced by discretization of objects in the image into the plurality of bins. Alternatively, or additionally, the image histogram can be a graphical representation of the number of pixels in an image as a function of their intensity. The bins are marked by a range of roll angles and indicate of one or more objects present in the image. Alternatively, or additionally, for determining orientation for a video, system may use a different number of bins for constructing the histogram, or a different function can be used to assign each object to a histogram bin.

The system determines 108 whether more than one object is present in the image or frame. If yes, the system counts 110 the number of objects in each of the bins. The bins may have any number of objects. In an example, the histogram has 4 bins B1, B2, B3, and B4 containing 4, 6, 7, and 26 objects, respectively.

The system then determines 112 highest number of objects ( $M$ ) in a single bin from all the bins. When the system determines the highest of count of objects in a single bin, it proceeds to apply an algorithm to calculate a winning bin in order to determine the orientation of the image. For example, when the system determines the highest number of objects ( $M$ ) in a single bin, the system calculates the value for the the expression " $M/2$ ". For example, the system determines that bin B4 has a maximum number of objects  $M=26$  and calculates  $M/2$ .

The system determines 114 whether there is only one bin with a number of objects greater than " $M/2$ ". For example, on calculating  $M/2 = 13$ , the system determines whether any of the other bins B1, B2, or B3 have a number of objects greater than 13. This algorithm is applied to determine the winning bin.

If there is only one bin with a number of objects greater than  $M/2$ , the system determines 116 the orientation of the image based on spatial orientation parameters for the winning bin which has a number of objects greater than  $M/2$ . The spatial orientation, i.e, the roll angle, Pan angle, tilt, etc. for the winning bin with a number of objects greater than  $M/2$  is selected to signify the correct orientation of the image when displayed.

The system updates 118 the orientation of the image or frame based on the orientation parameters of the winning bin. This updating can take the form of adding metadata, such as EXIF data, to the visual digital media file. The orientation parameters of the winning bin are expected to define the correct display orientation for the image or video frame. Thus, the system updates the orientation of the image or frame (or video fragment) according to the orientation parameters for the winning bin.

Alternatively, or additionally, for an image or a frame, there may be multiple bins with a number of objects greater than  $M/2$ . Such a scenario signifies detected objects having conflicting orientations in the image and thus no clear winner bin. In such a scenario, the system may not change 118 the display orientation of the image.

Fig. 2 is a block diagram of an exemplary environment that shows components of a system for implementing the techniques described in this disclosure. The environment includes client devices 210, servers 230, and network 240. Network 240 connects client devices 210 to servers 230. Client device 210 is an electronic device. Client device 210 may be capable of requesting and receiving data/communications over network 240. Example client devices 210 are personal computers (e.g., laptops), mobile communication devices, (e.g. smartphones, tablet computing devices), set-top boxes, game-consoles, embedded systems, an image capturing

device, a camera, a camcorder, video camera. Client device 210 may execute an application, such as a image capturing element 212 and a native image processing application 214. Server 230 may be a web server capable of sending, receiving and storing images or related metadata such as videos 232. Images or videos 232 may be stored on or accessible via server 230. Images or videos 232 may be associated with image processing application 214 and accessed using the application 214. When accessed, images or videos 232 may be transmitted and displayed on a client device, e.g., 210. Resources 218 and 218' are resources available to the client device 210 and/or applications thereon, or server(s) 230 and/or images or videos accessible therefrom, respectively. Resources 218' may be, for example, memory or storage resources; a text, image, video, audio, JavaScript, CSS, or other file or object; or other relevant resources. Network 240 may be any network or combination of networks that can carry data communication.

The subject matter described herein can be implemented in software and/or hardware (for example, computers, circuits, or processors). The subject matter can be implemented on a single device or across multiple devices (for example, a client device and a server device). Devices implementing the subject matter can be connected through a wired and/or wireless network. Such devices can receive inputs from a user (for example, from a mouse, keyboard, or touchscreen) and produce an output to a user (for example, through a display and/or a speaker). Specific examples disclosed are provided for illustrative purposes and do not limit the scope of the disclosure.

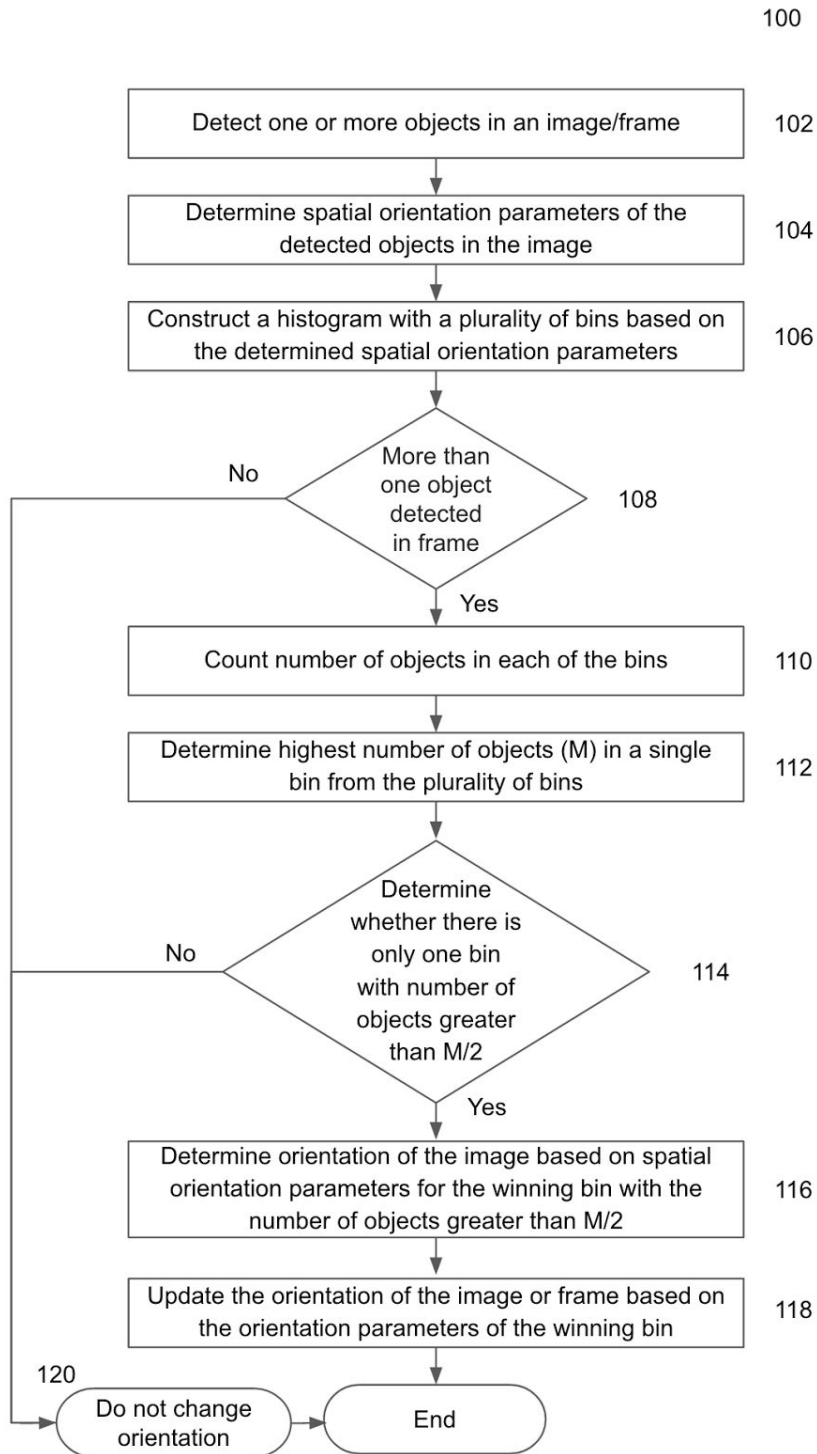
DRAWINGS

Fig. 1



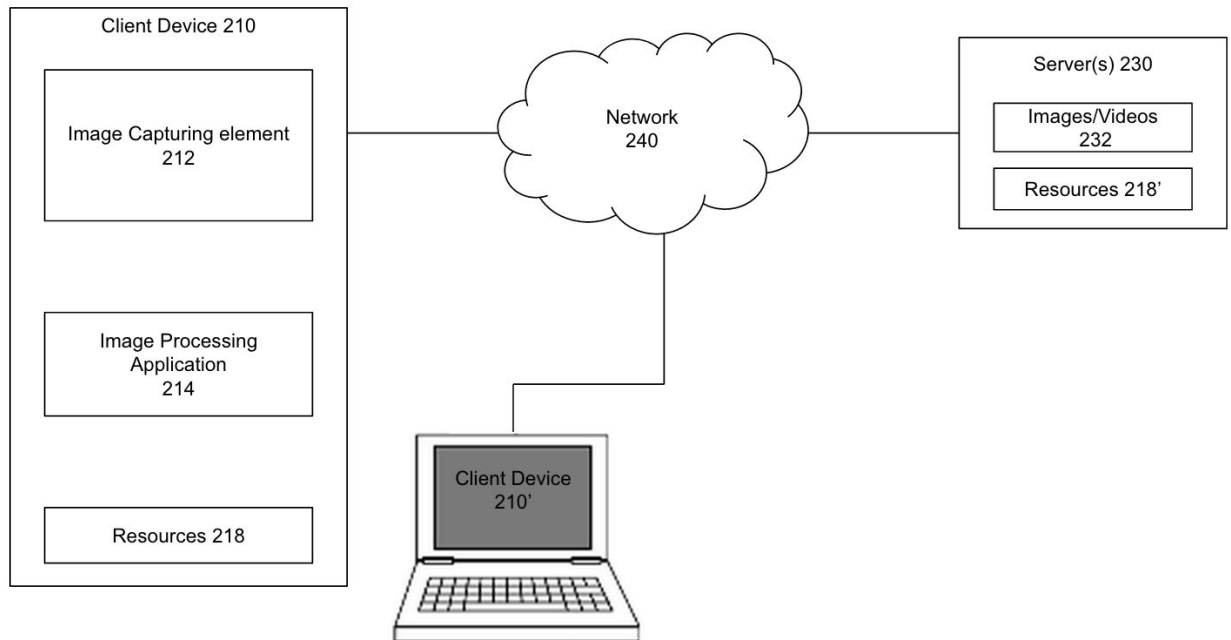


Fig. 2