AN IMAGE CAPTURE SYSTEM AND METHOD

FIELD

The present invention relates to a method for capturing images and an image capture system.

BACKGROUND

Most cameras are equipped with anti-shaking or image stabilizing functions (CCD movement) to prevent the camera movement effects for image/video shooting. This is generally indicated in low light conditions without the flash or when the flash is not sufficiently bright for the scene.

The Sony Cybershot Camera has a function called Panoramic capturing where a stationery user shoots a scenario by swinging the camera. This function stitches the images together to form a panorama image, but does not take into account the camera orientation or movement of the user. Moreover, this function cannot shoot a stationary subject from a large range of angles of view points.

There is recent interest in capturing of 3D images. One adopted format is stereoscopic capturing, while another is all around capture, which might be used in a volumetric display or to allow the user to rotate the view they wish to see at any angle. To capture all around images, one method is to put the subject(s) on a rotating table. But this method is limited by the object size and cannot be used for general purpose all round view shooting. Another method is

to use multiple cameras around the subject with interpolation or 3D modelling to generate the all round views. However, the generated views are not photo realistic and are not easily reusable.

SUMMARY

In general terms, the present invention relates to an inward shooting mechanism for capturing multiple view angles (0 to 360 degree, including 360 degree all-round views) of single/multiple subjects. The mechanism may include an image capturing device (camera) and a position/rotation measurement device (such as an Inertial Measurement Unit (IMU) sensor), which is moved by a consumer around a shooting subject guided by the information given by the IMU sensor. For example, the information may relate to providing feedback that the camera is too close, tilted, rotated and etc. The captured image data may then be corrected using image matching and camera orientation data from the IMU sensor. The combination of better quality input images and the use of image correction, which takes into account the camera orientation, may lead to generation of improved images.

According to a first aspect of the invention, there is provided a method of capturing images comprising determining an ideal track about a subject, calculating deviation information from the track during movement of a user, displaying feedback to the user depending on the deviation information, and capturing images of the subject during the movement of the user substantially about the track.

According to a second aspect of the invention, there is provided a system for image capture comprising an image sensor, an Inertial Measurement Unit (IMU), a display; and a processor configured to control the display to provide feedback to a user in relation to deviation from a determined track around a subject based on data from the IMU.

BRIEF DESCRIPTION OF THE DRAWINGS

Embodiments of the invention are disclosed hereinafter with reference to the accompanying drawings, in which:

Figure 1 is an illustration of an image capturing process using a camera, according to an example embodiment;

Figure 2 is a flow chart of an image processing method according to the example embodiment;

Figure 3 is a comparison between adopting unguided and guided user tracks;

Figure 4 is a graph of the Cartesian Coordinate Systems used in the calculations and algorithms;

Figure 5 is flow chart of an Inertial Measurement Unit (IMU) processing algorithm, according to the example embodiment;

Figure 6 is a schematic diagram of the hardware of the camera of Figure 1;

Figure 7 is a flow chart of a method of guiding the user to stay on the ideal path and orientation, according to the example embodiment;

Figure 8 is a screen shoot of the feedback displayed on the screen of the camera of Figure 1, which provides guidance to the user for staying on the ideal path and orientation;

Figure 9 is a flow chart of a feature matching algorithm utilised by the image processing method of Figure 2, according to the example embodiment;

Figure 10 is a flow chart of an algorithm, according to the example embodiment, which combines the results of the IMU processing algorithm of Figure 5, and feature matching algorithm of Figure 9;

Figure 11 is a flow chart of an image correction algorithm, according to the example embodiment; and

Figure 12 illustrates a method for calculating an ideal path around a subject using an autofocus sensor of the camera of Figure 1.

DETAILED DESCRIPTION

As depicted in Figure 1, a user holds a camera 102 and slowly walks around a subject 104 while keeping the camera 102 equidistant from, and accurately aimed at, the centre of the subject 104. As shown in Figure 3(a), the movement of the user will typically vary substantially from an ideal track 300 of moving the camera 102, in terms of the distance and camera orientation. According to an embodiment of the invention, Figure 3(b) shows a user guidance system 302 displayed on the screen (i.e. display unit) 304 of the camera 102, which is conveniently located on the back surface of the camera 102. In this way, the user is provided with guidance to maintain the camera 102 on the ideal track 300 with the correct orientation. Accordingly, Figure 3(c) shows that with the user guidance system 302, the ability of the user to maintain the camera 102 on the ideal track 300 for the case of the guided track is significantly improved over the case of the unguided track.

The camera 102 includes an Inertial Measurement Unit (hereinafter IMU) sensor 602. The IMU sensor 602 captures information from a gyroscope (not shown) and an accelerometer (not shown), and optionally an autofocus sensor 604 or a distance sensor (not shown). The data from the two sensors are combined to obtain pitch, yaw, roll, and movement parameters with reference to the [XYZ] Cartesian Coordinate System. The autofocus sensor 604 provides distance estimation to the subject 104 using infrared or sound waves. Alternatively the above mentioned data may be obtained using other suitable or equivalent sensor systems/ technologies.

The camera 102 is equipped with the display unit 304 to display the guidance feedback generated by the user guidance system 302, and installed with a memory (not shown) and a computation unit (i.e. processor) 606 to execute the software, which will be described in later parts of the specification. Figure 6 shows the camera 102, which includes the display unit 304, an image sensor 608, the autofocus sensor 604 and the IMU sensor 602, all being electrically coupled to the computation unit 606. The image sensor 608 may be a CCD/CMOS sensor, and the processor 606 may be a CPU or an application-specific integrated circuit (ASIC).

IMU

The IMU sensor 602 measures six-degrees-of-freedom, included a triple-axis gyroscope and a triple-axis accelerometer. To obtain better acceleration reading resolution, a measurement range of up to 4G may be sufficient. Additionally, to obtain better acceleration reading resolution, a measurement range up to 90

degrees may be sufficient. For example, an IMU sensor 602 made by Sparkfun Electronics, Model No.: 9DOF Razor IMU, AHRS compatible (but not installed with a magnetometer), and which uses the ITG-3200 triple-axis gyroscope, ADXL345 triple-axis accelerometer, and HMC 5883L triple-axis accelerometer (the deviation of 16G is not necessary) may be used.

User Interface

As shown in Figure 8, the display unit 304 of the camera 102 provides a number of different indications for feedback to the user about distance from the subject 104, and camera orientation. For example, a plurality of six axis arrows 802 gives feedback on yaw, roll and pitch. The colour of each arrow 802 indicates the level of error, so that when the user performs movement correction by adjusting closer to the ideal track 300, the colour changes towards green, while if the correction conversely deviates from the ideal track 300, the colour then switches towards red. Further, there is also a dotted box 804 representing the ideal camera position and a solid box 806 representing the actual camera position at that instance of use. Therefore, the user's objective is to coincide the two different boxes 804, 806 by repositioning the camera 102 through adjustments, such that the camera 102 is moved closer or further away from the subject 104 (in combinatory usage with the orientation correction described above). In the example shown in Figure 8, the user needs to rotate camera 102 clockwise and move camera 102 left, down and far away from the subject 104.

The User Interface displayed by the camera 102 may also provide a list of options prior to image capture, and/or another list of post processing options

after the image capture.

Software Algorithms

As shown in Figure 2, the camera 102 executes two main algorithms 200 during image capture. The first algorithm 200 is the image guidance system, and the second algorithm 200 being the image correction system. Specifically, the image guidance algorithm 200 is executed when the image capture function is activated. The image correction algorithm 200 is then executed for post processing either immediately after the image capture is completed, or subsequently on another separate device such as a computer. In either case, the IMU data generated by the IMU sensor 602 through usage of the user guidance system 302 are saved together with the uncorrected images.

Image Guidance Algorithm

With reference to Figure 7, the main steps for the image guidance algorithm 200 are:

1. An initialization step 702: Based on the initial position/rotation of the camera 102, and the distance information, a target camera movement path is predictively estimated;

2. An information collect step 704: Collection of the displacement and rotation information of the camera 102 from the IMU sensor 602; and

3. An adjustment step 706: Deciding the rotation and position adjustment for the camera 102 based on the collated IMU data, and display the corresponding information on the display unit 304 of the camera 102 for the

user to perform the necessary correction adjustments.

It is to be appreciated that both Steps 2 and 3 will be looped continuously, until conclusion of the image capturing.

Initial stage - Calculation of ideal path

As shown in Figure 12, the camera's 102 autofocus sensor 604 measures the distance of the camera 102 to the subject 104. Ideally, the centre of the subject 104 needs to be first located. Notably, the error percentage (between the estimated and ideal tracks) is larger if the camera 102 is very near to the subject 104. In this case, the user's movement should therefore be constrained so that he walks a smaller arc, because of the difficulty in estimating the ideal track 300. When the distance of the subject 104 to the camera 102 is sufficiently large, the user can then walk a full 360 degrees around the subject 104, since the error percentage is now considerably smaller. A bounding box may be displayed on the display unit 304 for informing the user the maximum size of the arc he needs to walk, in order to minimize the difference between the estimated and the ideal tracks. Alternatively, the user can also manually input the arc size and the radius of the track, if desired.

As shown in Figure 4, as the World Coordinate System (in relation to the [XYZ] planes) is defined at the initialisation stage, the camera 102 first needs to be positioned at a water level, measured with respect to a horizontal plane (it is to be appreciated that a deviation of approximately 10 degrees is an acceptable error tolerance range). To facilitate the performance of this step, suitable

related measurement sensors (e.g. a tilt sensor or a water level sensor) may be utilised to confirm the actual water level at which the camera 102 is positioned. Further, in this initial position, the velocity of the camera 102 is equal to zero, as the user has not begun moving around the subject 104. Consequently, a set of initial parameters (e.g. velocity, location coordinates, and etc) to provide reference basis for subsequent calculations along the track is obtained.

Information Collection IMU Calculations

To obtain the displacement information, the following steps shown in Figure 2 are used:

- Initializing, camera steady and water levelness to define a starting point for the track and the relevant coordinates;
- Integrating the gyroscope data to get the coordinate rotation;
- Removing the gravity component "g" from the accelerometer readings based on the coordinate rotation to obtain the real acceleration relative to the IMU sensor 602;
- Projecting the acceleration relative to the IMU sensor 602 back to the World Coordinate System; and
- Double integrating the acceleration to get the displacement in world coordinates.

These steps are shown in more detail in Figure 5. In particular, the IMU sensor 602 provides the following output signals $R_x, R_y, R_z, a_x, a_y, a_z, t$, where "*R*" represents the rotation speed (as degrees per second) of the axis and "*a*"

represents the acceleration (as meters per square second) in the IMU coordinate system previously shown in Figure 4.

The IMU coordinate system is defined with reference to the [XYZ] planes and the actual rotation in the IMU system (as depicted in Figure 4) is denoted as α, β, γ, where the output of the gyroscope is integrated over time to get the incremental coordinate transformation in the IMU coordinate system, which is expressed as the group of equations (1):

$$\begin{aligned} \alpha &= R_x \times t \\ \beta &= R_y \times t \\ \gamma &= R_z \times t \end{aligned} \qquad (1)$$

The initial value is then obtained from the acceleration reading of the camera 102 when the camera 102 is in a stationary position, the equations used for obtaining the initial value being expressed as equations (2) to (4):

$$a_0 = [a_{x0}, a_{y0}, a_{z0}]' \qquad (2)$$

$$\begin{aligned} \alpha_0 &= -a\sin(a_{y0}/g) \\ \beta_0 &= a\sin(a_{x0}/g) \\ \gamma_0 &= 0 \end{aligned} \qquad (3)$$

$$R_0 = \begin{bmatrix} \cos\gamma_0 & -\sin\gamma_0 & 0 \\ \sin\gamma_0 & \cos\gamma_0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \bullet \begin{bmatrix} \cos\beta_0 & 0 & \sin\beta_0 \\ 0 & 1 & 0 \\ -\sin\beta_0 & 0 & \cos\beta_0 \end{bmatrix} \bullet \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha_0 & -\sin\alpha_0 \\ 0 & \sin\alpha_0 & \cos\alpha_0 \end{bmatrix} \qquad (4)$$

The accumulated IMU coordinate at time *t* hence would be expressed as equation (5):

$$R_t = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix} \bullet \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix} \bullet \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix} \bullet R_{t-1} \qquad (5)$$

To remove the gravity component, "g", from the accelerometer readings, it is first necessary to transform them into their respective rotational components by inversing the Rotation matrix "R_t" of equation (5).

Further, the Initial Gravity Vector is expressed in equation (6) as:

$$G_0 = [0,0,-g] \qquad (6)$$

Additionally, the Gravity vector when converted to the Time-domain, "t", is expressed as equation (7):

$$G_t = inverse(R_t) \bullet G_0 \qquad (7)$$

Thus the real acceleration vector expressed in the IMU coordinate system is defined as equation (8):

$$a_t = [a_x, a_y, a_z]' - G_t \qquad (8)$$

Following on, this allows the real acceleration vector in the world coordinate to be calculated based on the formula expressed in equation (9):

$$A_t = R_t \bullet a_t \qquad (9)$$

Subsequently, this is integrated to provide the velocity vectors in the world coordinate as expressed in equations (10):

$$\begin{aligned} V_t &= V_{t-1} + A_t \times t \\ V_0 &= 0 \end{aligned} \qquad (10)$$

The equations (10) are then consequently integrated to provide the displacement vectors in the world coordinate which are expressed as equations (11):

$$\begin{aligned} D_t &= D_{t-1} + (V_t + V_{t-1}) \times t/2 \\ D_0 &= 0 \end{aligned} \qquad (11)$$

Using Kalman Filter and Low Pass Filter

In addition, conventional signal processing technology such as Kalman filter and Finite Impulse Response (FIR) filter, both of which are configured as low-pass filters to be used in the invention, may be adopted to compensate for the gyroscopic drift/noise, accelerometric noise and velocity drift shown in Figure 5. The Kalman filter and/or moving average algorithm may be used to compensate mainly for drift that is based on previously obtained results to predict future movement. On the other hand, the low-pass FIR filter may be used to remove high frequency noises, or alternatively provide band pass filtering functions.

Adjustment Step - Error From Ideal Path – Generate Feedback Images

The information collected is then compared to the ideal track 300. Based on the error values or deviation, the appropriate rotation and position adjustment to be effected for the camera 102, are shown on the display unit 304 to guide the user to perform the corresponding required movement adjustments.

Image Correction

The image correction system 200 comprises two sub-algorithms. The first sub-algorithm is based on image features matching (e.g. Scale-invariant feature transform (SIFT)) which gives an effect similar to averaging the images. The second sub-algorithm relates to performing corrections based on the actual camera orientation from the IMU sensor 602. The camera shift, rotation and scaling away from the track may then be calculated by combining the image features matching results and the IMU output. The captured image is corrected based on a three-by-three matrix which contains the relevant shift, rotate and

scaling information. The final output then correspond to the corrected images of the all round views.

Image Matching

As shown in Figure 9, in relation to two continuous images, the matching process identifies features points for each image based on an image feature extraction algorithm such as SIFT, Affine SIFT (ASIFT), the Harris affine region detector, or the Hessian affine region detector. When each feature point has been detected, a corresponding feature descriptor (which is typically a vector identifier) is then associated to it (e.g. with SIFT usage, a vector with 128 elements is generated). The feature points from the two images are then compared and matched based on the similarity of the feature descriptors and other miscellaneous parameter constraints (e.g. location). The matched result is eventually used to determine how much the camera 102 has shifted and rotated in terms of the pitch direction.

Combining Image Matching and IMU Data

As shown in Figure 10, the IMU data previously collected by the image guidance algorithm 200 and image matching results are combined according to certain defined weighting constants, as shown in the list below, to provide a set of variables used by the image correction algorithm 200. The weight constants are:

1. Pitch, roll, and yaw variables (weighted mainly to the IMU sensor 602, but with some degree to image matching);

2. Image shift (weighted mainly to image matching); and

3. Scale (weighted mainly to the IMU sensor 602).

Compensation

As shown in Figure 11, there are three primary steps involved in performing image correcting, which includes rotation, rescaling and shifting compensation.

(1) Rotation compensation:

From the integrated pitch, roll, and yaw variables, $[\varphi \quad \phi \quad \theta]$, obtained for each frame, and together with the camera's configured internal parameters, which are specified in the form of a "*K*" matrix (which is a three-by-three matrix and contains parameters related to the camera hardware configuration), the Rotation Matrix, "R_2", is calculated as follows in equation (12):

$$R_2 = \begin{bmatrix} \cos\theta\cos\varphi & -\cos\phi\sin\varphi + \sin\phi\sin\theta\cos\varphi & \sin\phi\sin\varphi + \cos\phi\sin\theta\cos\varphi \\ \cos\theta\sin\varphi & \cos\phi\cos\varphi + \sin\phi\sin\theta\sin\varphi & -\sin\phi\cos\varphi + \cos\phi\sin\theta\sin\varphi \\ -\sin\theta & \sin\phi\cos\theta & \cos\phi\cos\theta \end{bmatrix} \quad (12)$$

This is then applied to each image according to the Compensated Image, which is expressed in equation (13) as:

$$I_1 = KRK^{-} \bullet I_0 \quad (13)$$

(2) Scaling compensation:

The Scaling Factor, "*S*", is calculated according to the distance to the subject 104 from the IMU data for each image or the ideal distance from the initialisation step. This is subsequently applied to each rotated image according to the Compensated Image equation (14) specified as:

$$I_2 = S \times I_1 \quad (14)$$

(3) Shift compensation:

The Shift Value, $A = \begin{bmatrix} 1 & 0 & X \\ 0 & 1 & Y \\ 0 & 0 & 1 \end{bmatrix}$, is calculated based on the image shift variables obtained from the image matching calculations. This is subsequently applied to each scaled image according to the Final Compensated Image equation (15) expressed as:

$$I = A \bullet I_2 \qquad (15)$$

Eliminating Extra Frames

When the user moves along the track, the speed of movement is likely to vary and the corresponding rotation around the object 104 may also fluctuate between fast or slow, or may even stop occasionally. Therefore, to obtain a more constant frame rate along the track, the user may choose an option to eliminate any extra frames generated, when the camera 102 moves slowly on the track, around the object 104.

The foregoing described embodiments have the following advantages over known conventional systems:

- Ease of allowing the user to create an all around image sequence or video for any subject;
- Not giving rise to undesirable image artifacts due to generation of virtual images using methods such as interpolation, text mapping, photo synthesis, and etc;

- Provision of a relatively simple method, which guides the user to follow a pre-designed path surrounding the subject of interest;
- Allowing the user to create multi-perspective contents for user consumer-grade imaging devices, in which these contents may in turn be used for multi-perspective 3D displays; and
- Providing the ability to easily adapt the same generated contents for viewing in 2D displays by simply rotating the contents.

While the invention has been illustrated and described in detail in the drawings and foregoing description, such illustration and description are to be considered illustrative or exemplary, and not restrictive; the invention is not limited to the disclosed embodiments. Other variations to the disclosed embodiments can be understood and effected by those skilled in the art in practising the claimed invention.

Claims

1. A method of capturing images comprising:

determining an ideal track about a subject;

calculating deviation information from the track during movement of a user;

displaying feedback to the user depending on the deviation information; and

capturing images of the subject during the movement of the user substantially about the track.

2. The method in claim 1 wherein determining the track comprises determining the distance to the subject, inputting user desired capture variables and calculating the track depending on the distance and variables.

3. The method in claim 3 wherein determining the distance comprises guiding the user to hold a camera steady, level, aimed at the centre of the subject and/or the correct scale relative to the level of zoom.

4. The method in claim 3 wherein determining the distance is based on a signal from an autofocus or distance sensor.

5. The method in any preceding claim wherein the deviation is calculated based on acceleration and/or orientations signals from an Inertial Measurement Unit (IMU).

6. The method in claim 5 further comprising filtering the IMU signals.

7. The method in claim 6 further comprising removing the gravity component from the filtered IMU signals.

8. The method in claim 7 further comprising integrating the IMU signals to convert to a world coordinate system.

9. The method in any preceding claim further comprising correcting the captured images for rotation, scale and shift.

10. The method in claim 9 wherein the correcting is based on the deviation information.

11. The method in claim 10 further comprising image matching the captured images, wherein the correcting is also based on image matching information.

12. The method in claim 11 wherein the correcting is based on a weighting between imaging matching information and the deviation information.

13. The method in claim 11 or 12 when dependant on any of claims 5 to 8, wherein the deviation is also calculated based on the image matching information.

14. The method in any preceding claim wherein the feedback comprises displaying graphical images indicating how the user should correct the location or orientation of the camera.

15. A system for image capture comprising:

an image sensor;

an Inertial Measurement Unit (IMU);

a display; and

a processor configured to control the display to provide feedback to a user in relation to deviation from a determined track around a subject based on data from the IMU.

16. The system in claim 15 wherein the processor is further configured to correct image captured by the image sensor based on data from the IMU.

17. The system in claim 15 further comprising an autofocus sensor, and the processor is further configured to determine the track based on the distance to the subject determined from the autofocus sensor.