



US 20170078649A1

(19) **United States**

(12) **Patent Application Publication**
Taubin et al.

(10) **Pub. No.: US 2017/0078649 A1**

(43) **Pub. Date: Mar. 16, 2017**

(54) **METHOD AND SYSTEM FOR UNSYNCHRONIZED STRUCTURED LIGHTING**

Publication Classification

(51) **Int. Cl.**
H04N 13/02 (2006.01)
G06T 7/00 (2006.01)
H04N 13/00 (2006.01)
(52) **U.S. Cl.**
CPC *H04N 13/0253* (2013.01); *H04N 13/0051* (2013.01); *G06T 7/0057* (2013.01); *G06T 2207/10028* (2013.01)

(71) Applicant: **Brown University**, Providence, RI (US)

(72) Inventors: **Gabriel Taubin**, Providence, RI (US);
Daniel Moreno, Norwich, CT (US);
Fatih Calakli, Providence, RI (US)

(21) Appl. No.: **15/124,176**

(22) PCT Filed: **Mar. 9, 2015**

(86) PCT No.: **PCT/US15/19357**

§ 371 (c)(1),

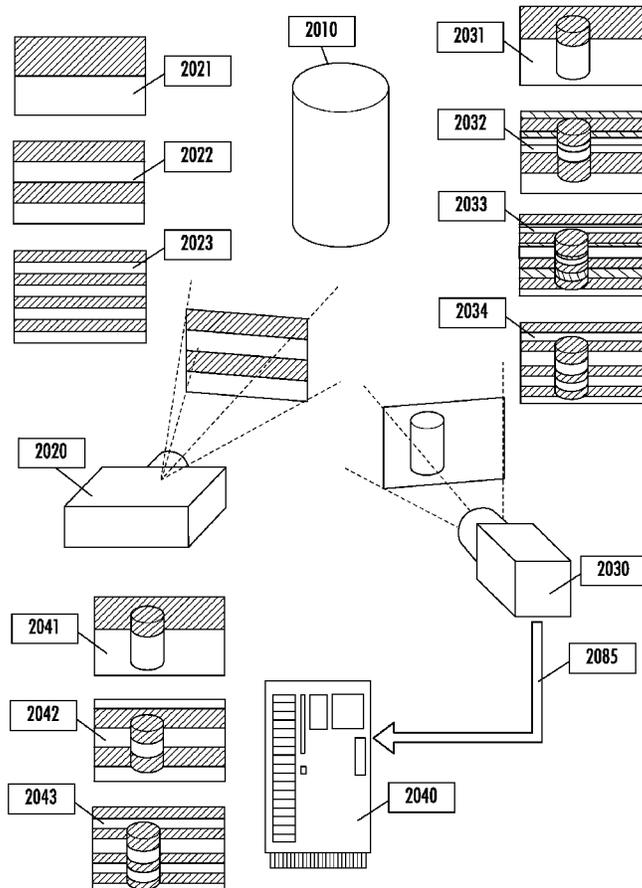
(2) Date: **Sep. 7, 2016**

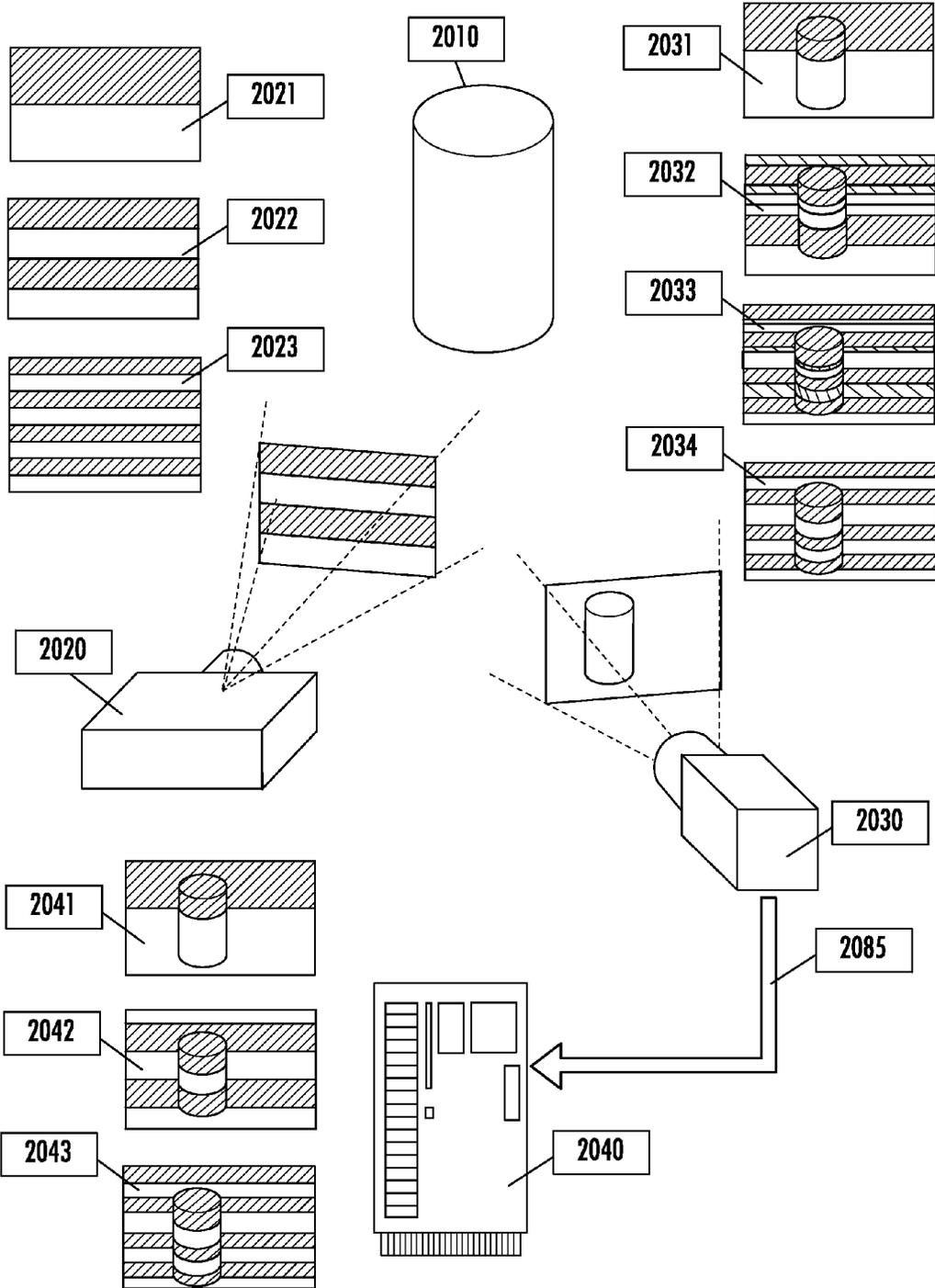
(57) **ABSTRACT**

A system and method to capture the surface geometry a three-dimensional object in a scene using unsynchronized structured lighting is disclosed. The method and system includes a pattern projector configured and arranged to project a sequence of image patterns onto the scene at a pattern frame rate, a camera configured and arranged to capture a sequence of unsynchronized image patterns of the scene at an image capture rate, and a processor configured and arranged to synthesize a sequence of synchronized image frames from the unsynchronized image patterns of the scene. Each of the synchronized image frames corresponds to one image pattern of the sequence of image patterns.

Related U.S. Application Data

(60) Provisional application No. 61/949,529, filed on Mar. 7, 2014.





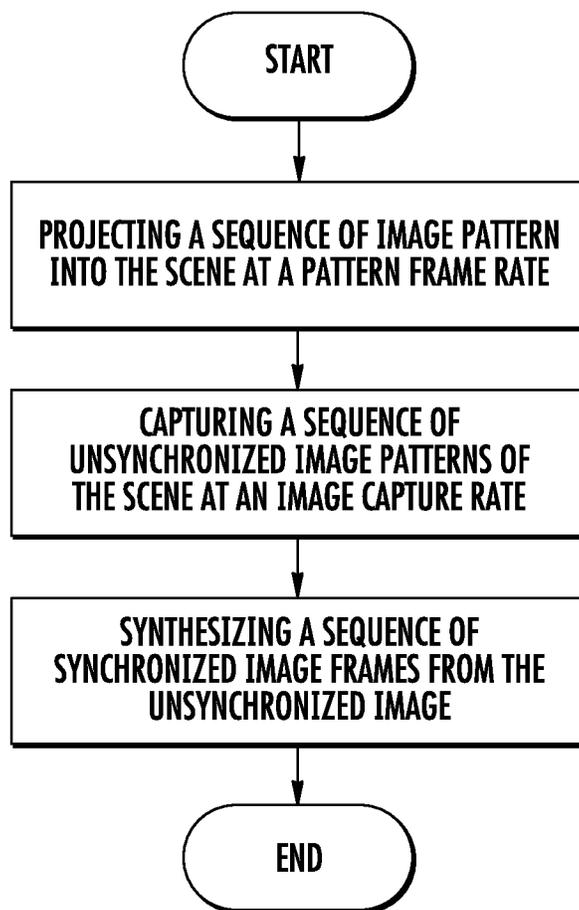


FIG. 2

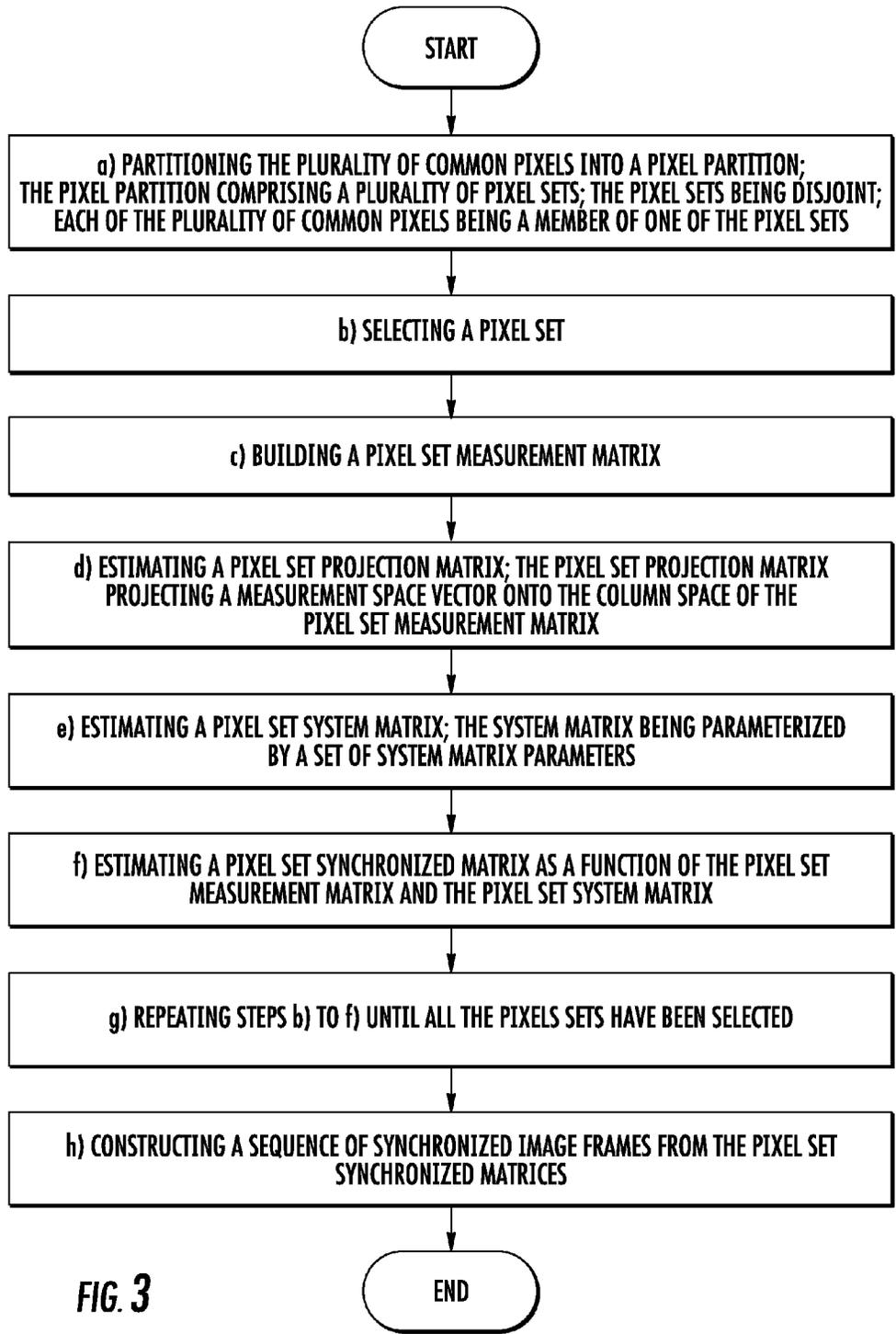


FIG. 3

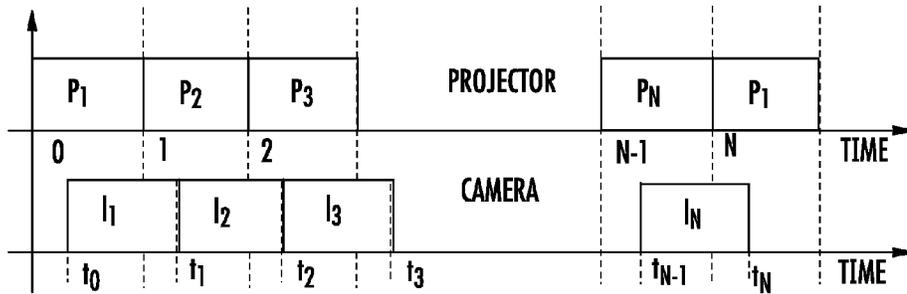


FIG. 4

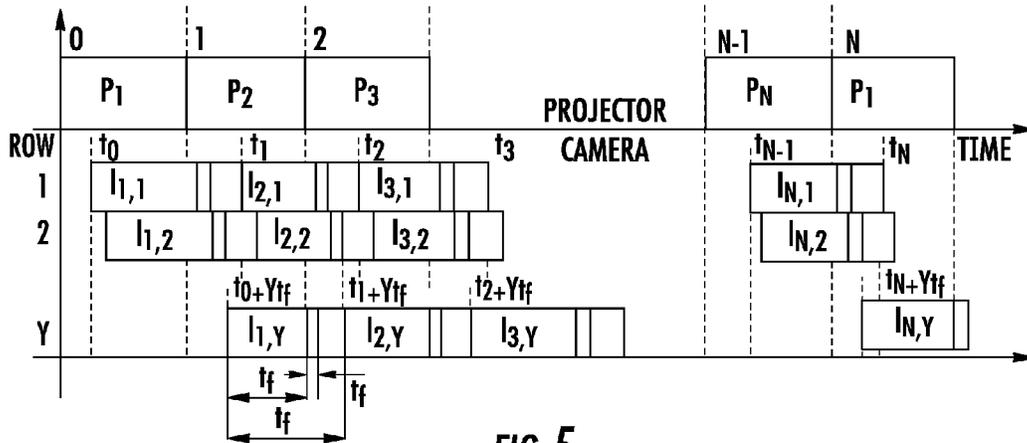


FIG. 5

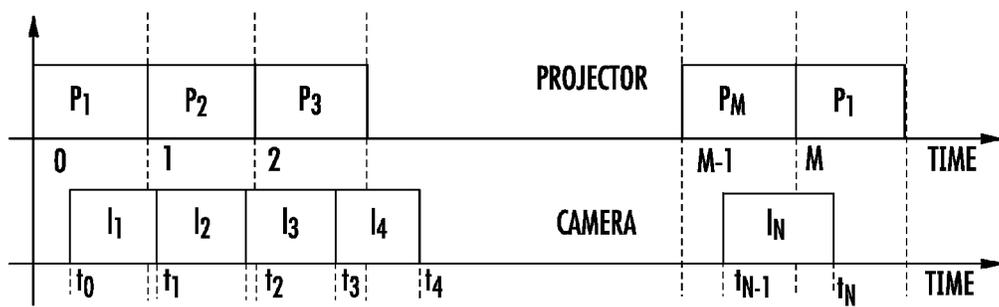


FIG. 6

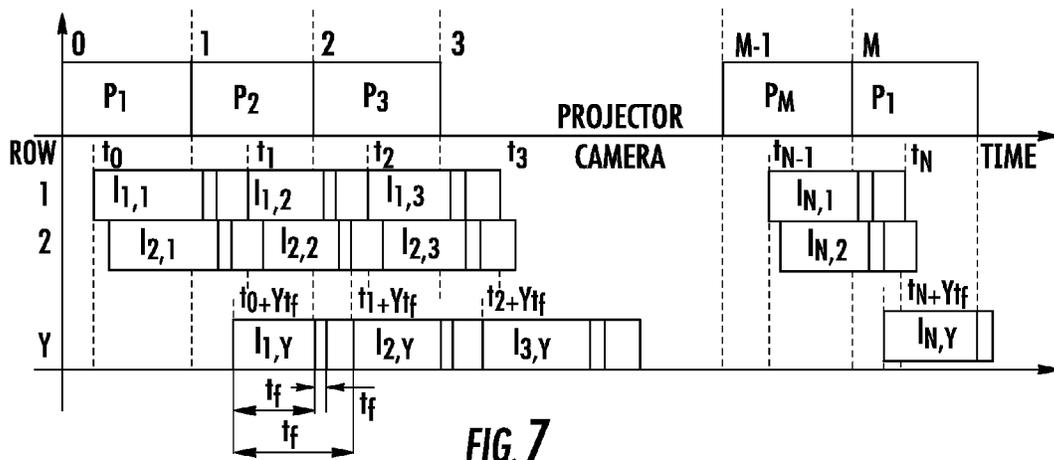


FIG. 7

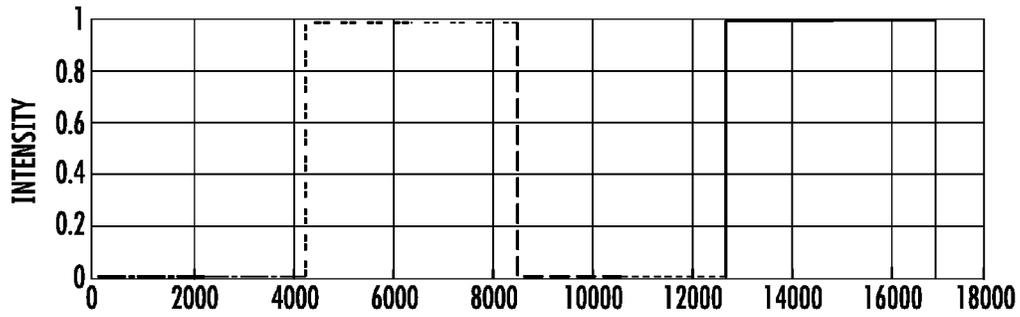


FIG. 8

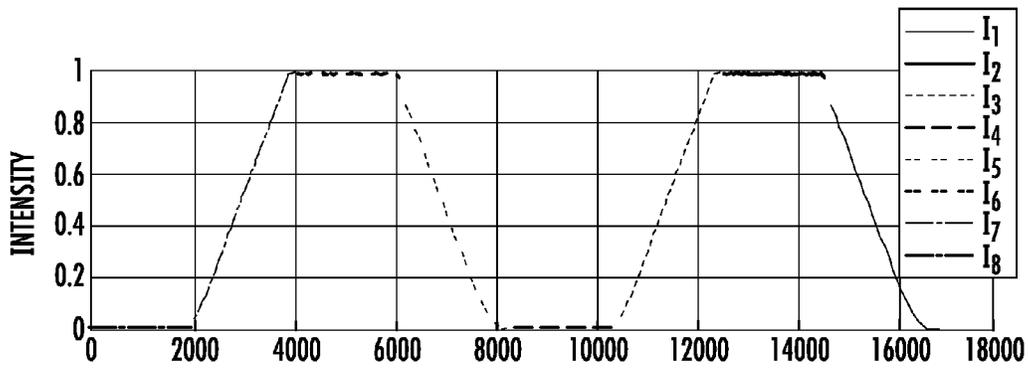


FIG. 9

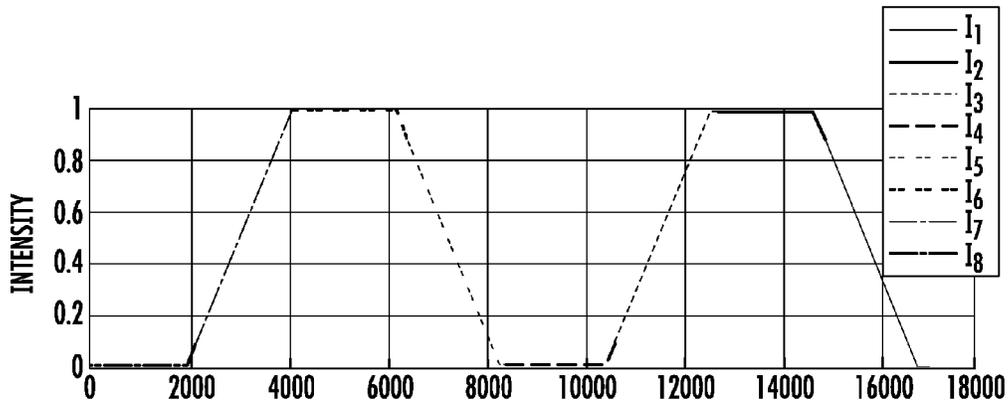


FIG. 10

METHOD AND SYSTEM FOR UNSYNCHRONIZED STRUCTURED LIGHTING

CROSS-REFERENCE TO RELATED APPLICATION

[0001] This application claims priority to earlier filed U.S. Provisional Application Ser. No. 61/949,529, filed Mar. 7, 2014, the contents of which are incorporated herein by reference.

BACKGROUND OF THE INVENTION

[0002] 1. Field of the Invention

[0003] This patent document relates generally to the field of three-dimensional shape capture of the surface geometry of an object, and more particularly to structured lighting three-dimensional shape capture.

[0004] 2. Background of the Related Art

[0005] Three-dimensional scanning and digitization of the surface geometry of objects is commonly used in many industries and services, and their applications are numerous. A few examples of such applications are inspection and measurement of shape conformity in industrial production systems, digitization of clay models for industrial design and styling applications, reverse engineering of existing parts with complex geometry for three-dimensional printing, interactive visualization of objects in multimedia applications, three-dimensional documentation of artwork, historic and archaeological artifacts, human body scanning for better orthotics adaptation, biometry or custom-fit clothing, and three-dimensional forensic reconstruction of crime scenes.

[0006] One technology for three-dimensional shape capture is based on structured lighting. Three dimensional shape capture systems based on structure lighting are more accurate than those based on time-of-flight (TOF) image sensors. In a standard structured lighting 3D shape capture system a pattern projector is used to illuminate the scene of interest with a sequence of known two-dimensional patterns, and a camera is used to capture a sequence of images, synchronized with the projected patterns. The camera captures one image for each projected pattern. Each sequence of images captured by the camera is decoded by a computer processor into a dense set of projector-camera pixel correspondences, and subsequently into a three-dimensional range image, using the principles of optical triangulation.

[0007] The main limitation of three-dimensional shape capture systems is the required synchronization between projector and camera. To capture a three-dimensional snapshot of a moving scene, the sequence of patterns must be projected at a fast rate, the camera must capture image frames exactly at the same frame rate, and the camera has to start capturing the first frame of the sequence exactly when the projector starts to project the first pattern.

[0008] Therefore, there is a need for three-dimensional shape measurement methods and systems based on structure lighting where the camera and the pattern projector are not synchronized.

[0009] Further complicating matters, image sensors generally use one of two different technologies to capture an image, referred to as “rolling shutter” and “global shutter”. “Rolling shutter” is a method of image capture in which a still picture or each frame of a video is captured not by taking a snapshot of the entire scene at single instant in time

but rather by scanning across the scene rapidly, either vertically or horizontally. In other words, not all parts of the image of the scene are recorded at exactly the same instant. This is in contrast with “global shutter” in which the entire frame is captured at the same instant. Even though most image sensors in consumer devices are rolling shutter sensors, many image sensors used in industrial applications are global shutter sensors.

[0010] Therefore, there is a need for three-dimensional shape measurement methods and systems based on structure lighting where the camera and the pattern projector are not synchronized, supporting both global shutter and rolling shutter image sensors.

SUMMARY OF THE INVENTION

[0011] A system and method to capture the surface geometry a three-dimensional object in a scene using unsynchronized structured lighting solves the problems of the prior art. The method and system includes a pattern projector configured and arranged to project a sequence of image patterns onto the scene at a pattern frame rate, a camera configured and arranged to capture a sequence of unsynchronized image patterns of the scene at an image capture rate; and a processor configured and arranged to synthesize a sequence of synchronized image frames from the unsynchronized image patterns of the scene, each of the synchronized image frames corresponding to one image pattern of the sequence of image patterns. Because the method enables use of an unsynchronized pattern projector and camera significant cost savings can be achieved. The method enables use of inexpensive cameras, such as smartphone cameras, webcams, point-and-shoot digital cameras, camcorders as well as industrial cameras. Furthermore, the method and system enable processing the images with a variety of computing hardware, such as computers, digital signal processors, smartphone processors and the like. Consequently, three-dimensional image capture using structured lighting may be used with relatively little capital investment.

BRIEF DESCRIPTION OF THE DRAWINGS

[0012] These and other features, aspects, and advantages of the method and system will become better understood with reference to the following description, appended claims, and accompanying drawings where:

[0013] FIG. 1 is an illustration of an exemplary embodiment of the method and system for unsynchronized structured lighting;

[0014] FIG. 2 shows a flowchart of an exemplary embodiment of the method and system for unsynchronized structured lighting;

[0015] FIG. 3 shows a flowchart of an exemplary embodiment of the method and system for unsynchronized structured lighting of synthesizing the synchronized sequence of image frames;

[0016] FIG. 4 is a chart illustrating the timing for a method to synthesize the synchronized sequence of image frames from an unsynchronized sequence of image frames, where a global shutter image sensor is used and where the image frame rate is identical to the pattern frame rate;

[0017] FIG. 5 is a chart illustrating the timing for a method to synthesize the synchronized sequence of image frames from an unsynchronized sequence of image frames, where a

rolling shutter image sensor is used and where the image frame rate is identical to the pattern frame rate;

[0018] FIG. 6 is a chart illustrating the timing for a method to synthesize the synchronized sequence of image frames from an unsynchronized sequence of image frames, where a global shutter image sensor is used and where the image frame rate is higher or equal than the pattern frame rate

[0019] FIG. 7 is a chart illustrating the timing for method to synthesize the synchronized sequence of image frames from an unsynchronized sequence of image frames, where a rolling shutter image sensor is used and where the image frame rate is higher or equal than the pattern frame rate;

[0020] FIG. 8 shows a chart illustrating a calibration pattern used to correct the time;

[0021] FIG. 9 shows a chart illustrating image data normalized according to the method described herein; and

[0022] FIG. 10 shows a chart illustrating a model estimating the pattern value for each pixel in a captured image.

DESCRIPTION OF THE PREFERRED EMBODIMENT

[0023] A system and method to capture the surface geometry a three-dimensional object in a scene using unsynchronized structured lighting is shown generally in FIGS. 1 and 2. The method and system includes a pattern projector configured and arranged to project a sequence of image patterns onto the scene at a pattern frame rate, a camera configured and arranged to capture a sequence of unsynchronized image patterns of the scene at an image capture rate; and a processor configured and arranged to synthesize a sequence of synchronized image frames from the unsynchronized image patterns of the scene, each of the synchronized image frames corresponding to one image pattern of the sequence of image patterns.

[0024] One object of the present invention is a system to synthesize a synchronized sequence of image frames from an unsynchronized sequence of image frames, illustrated in FIG. 1; the unsynchronized image frames captured while a three-dimensional scene was illuminated by a sequence of patterns. The system comprises a pattern projector **2020** and a camera **2030**. An object **2010** is partially illuminated by the pattern projector **2020**, and partially visible by the camera **2030**. The pattern projector projects a sequence of patterns **2021**, **2022**, **2023** at a certain pattern rate. The pattern rate is measured in patterns per second. The camera **2030** captures a sequence of unsynchronized image frames **2031**, **2032**, **2033**, **2034**, at a certain frame rate. The frame rate is larger or equal than the pattern rate. The number of unsynchronized image frames captured by the camera is larger or equal than the number of patterns in the sequence of patterns. The camera starts capturing the first unsynchronized image frame not earlier than the time when the projector starts projecting the first pattern of the sequence of image patterns. The camera ends capturing the last unsynchronized image frame not earlier than the time when the projector ends projecting the last pattern of the sequence of image patterns. To capture a frame the camera opens the camera aperture, which it closes after an exposure time. The camera determines the pixel values by integrating the incoming light while the aperture is open. Since camera and projector are not synchronized, the projector may switch patterns while the camera has the aperture open. As a result, some or all of the pixels of the captured unsynchronized image frame **2032** will be partially exposed to two consecu-

tive patterns **2021** and **2022**. The resulting sequence of unsynchronized image frames, are transmitted to a computer processor **2040**, which executes a method to synthesize a synchronized sequence of image frames from the unsynchronized sequence of image frames. The number of frames in the synchronized sequence of image frames will be the same as the number of patterns and represent estimates of what the camera would have captured if it were synchronized with the projector. In a preferred embodiment the camera **2030** and the computer processor **2040** are components of a single device such a digital camera, smartphone, or computer tablet.

[0025] Another object of the invention is an unsynchronized three-dimensional shape capture system, comprising the system to synthesize a synchronized sequence of image frames from an unsynchronized sequence of image frames described above, and further comprising prior art methods for decoding, three-dimensional triangulation, and optionally geometric processing, executed by the computer processor.

[0026] Another object of the invention is a three-dimensional snapshot camera comprising the unsynchronized three-dimensional shape capture system, where the projector has the means to select the pattern rate from a plurality of supported pattern rates, the camera has the means to select the frame rate from a plurality of supported frame rates, and the camera is capable of capturing the unsynchronized image frames in burst mode at a fast frame rate. In a preferred embodiment the projector has a knob to select the pattern rate. In another preferred embodiment the pattern rate is set by a pattern rate code sent to the projector through a communications link. Furthermore, the system has means to set the pattern rate and the frame rate so that the frame rate is not slower than the pattern rate. In a more preferred embodiment the user sets the pattern rate and the frame rate.

[0027] In a more preferred embodiment of the snapshot camera, the camera has the means to receive a camera trigger signal, and the means to set the number of burst mode frames. In an even more preferred embodiment, the camera trigger signal is generated by a camera trigger push-button. When the camera receives the trigger signal it starts capturing the unsynchronized image frames at the set frame rate, and it stops capturing unsynchronized image frames after capturing the set number of burst mode frames.

[0028] In a first preferred embodiment of the snapshot camera with camera trigger signal, the projector continuously projects the sequence of patterns in a cyclic fashion. In a more preferred embodiment the system has the means of detecting when the first pattern is about to be projected, and the camera trigger signal is delayed until that moment.

[0029] In a second preferred embodiment of the snapshot camera with camera trigger signal, the projector has the means to receive a projector trigger signal. In a more preferred embodiment the camera generates the projector trigger signal after receiving the camera trigger signal, and the camera has the means to send the projector trigger signal to the projector. In an even more preferred embodiment the camera has a flash trigger output, and it sends the projector trigger signal to the projector through the flash trigger output. When the projector receives the trigger signal it starts projecting the sequence of patterns at the set pattern rate, and it stops projecting patterns after it projects the last pattern.

[0030] Another object of this invention is a method to synthesize a synchronized sequence of image frames from an unsynchronized sequence of image frames, generating a number of frames in the synchronized sequence of image frames equal to the number of projected patterns, and representing estimates of what the camera would have captured if it were synchronized with the projector.

[0031] As will be described in greater detail below in the associated proofs, the

[0032] method to synthesize the synchronized sequence of image frames from the unsynchronized sequence of image frames is shown generally in FIG. 3. Each of the synchronized image frames corresponds to one of a sequence of image patterns. Further, the synchronized image frames and unsynchronized image frames have the same width and height. Each image frame includes a plurality of common pixels. In a first step a) the plurality of common pixels is partitioned into a pixel partition. The pixel partition includes a plurality of pixel sets, each of which is disjoint. Further each of the plurality of common pixels is a member of one of the pixel sets. In a step b) a pixel set is selected. In a step c) a pixel set measurement matrix is built. In a step d) a pixel set projection matrix is estimated. The pixel set projection matrix projects a measurement space vector onto the column space of the pixel set measurement matrix. In a step e) a pixel set system matrix is estimated. The system matrix is parameterized by a set of system matrix parameters. In a step f) a pixel set synchronized matrix as a function of the pixel set measurement matrix and the pixel set system matrix is estimated. In a step g) steps b) to f) are repeated until all the pixels sets have been selected. In a step h) a sequence of synchronized image frames from the pixel set synchronized matrices is constructed.

[0033] In a preferred embodiment, the method to synthesize the synchronized sequence of image frames from an unsynchronized sequence of image frames, applies to a global shutter image sensor where the image frame rate is identical to the pattern frame rate. FIG. 4 illustrates the timing for this embodiment. In this embodiment, the projector projects N patterns at a fixed frame rate, a global shutter image sensor capture N images at identical frame rate. Capturing each image takes exactly one unit of time, normalized by the projector frame rate. The start time for the first image capture t_0 is unknown, but the start time for the n-th image capture is related to the start time for the first image capture $t_n = t_0 + n - 1$. The actual value measured by the image sensor at the (x, y) pixel of the n-th image, can be modeled as

$$I_n(x, y) = (1 - t_0)P_n(x, y) + t_0P_{n+1}(x, y)$$

[0034] where $P_n(x, y)$ and $P_{n-1}(x, y)$ represent the pattern values to be estimated that contribute to the image pixel (x, y) and $P_{N+1} = P_1$. Projected patterns are known in advance, but since it is not known which projector pixel illuminates each image pixel, they have to be treated as unknown. To estimate the value of t_0 , the following expression is minimized

$$E(t_0) = \frac{1}{2} \sum_{n=1}^N \sum_{(x,y)} ((1 - t_0)P_n(x, y) + t_0P_{n+1}(x, y) - I_n(x, y))^2$$

[0035] with respect to t_0 , where the sum is over a subset of pixels (x, y) for which the corresponding pattern pixel values $P_n(x, y)$ and $P_{n-1}(x, y)$ are known. Differentiating $E(t_0)$ with respect to t_0 , and equating the result to zero, an expression to estimate t_0 is obtained

$$t_0 = \frac{\sum_{n=1}^N \sum_{(x,y)} (P_{n+1}(x, y) - P_n(x, y))(I_n(x, y) - P_n(x, y))}{\sum_{n=1}^N \sum_{(x,y)} (P_{n+1}(x, y) - P_n(x, y))^2}$$

[0036] Once the value of t_0 has been estimated, the N pattern pixel values $P_1(x, y), \dots, P_N(x, y)$ can be estimated for each pixel (x, y) by minimizing the following expression

$$E(P_1(x, y), \dots, P_N(x, y)) = \frac{1}{2} \sum_{n=1}^N ((1 - t_0)P_n(x, y) + t_0P_{n+1}(x, y) - I_n(x, y))^2$$

which reduces to solving the following system of N linear equations

$$\beta P_{n-1}(x, y) + \alpha P_n(x, y) + \beta P_{n+1}(x, y) = t_0 I_{n-1}(x, y) + (1 - t_0) I_n(x, y)$$

for $n=1, \dots, N$, where $\alpha = t_0^2 + (1 - t_0)^2$ and $\beta = t_0(1 - t_0)$.

[0037] In another preferred embodiment, the method to synthesize the synchronized sequence of image frames from an unsynchronized sequence of image frames, applies to a rolling shutter image sensor where the image frame rate is identical to the pattern frame rate. FIG. 5 illustrates the timing for this embodiment. We project N patterns at fixed framerate, a rolling shutter camera captures N images. Capture begins while pattern P_1 is being projected. Projector framerate is 1, pattern P_n is projected between time $n-1$ and n . A camera frame is read every t_r time, camera framerate is assumed equal to projector framerate but in practice may vary a little. A camera row requires t_r time to be readout from the sensor, thus, a sensor with Y rows needs a time $Y t_r$ to read a complete frame, $Y t_r \leq t_e$. Each camera frame is exposed t_e time, its readout begins immediately after exposure ends, $t_e + t_r \leq t_f$.

[0038] Camera row y in image n begins being exposed at time $t_{n,y}$

$$t_{n,y} = t_0 + (n-1)t_r + y t_r, y: 0 \dots Y-1,$$

[0039] and exposition ends at time $t_{n,y} + t_e$

[0040] In this model image n is exposed while pattern P_n and P_{n+1} are being projected. Intensity level measured at a pixel in row y is given by

$$I_{n,y} = (n - t_{n,y})k_{n,y}P_n + (t_{n,y} + t_e - n)k_{n,y}P_{n+1} + C_{n,y}$$

[0041] The constants $k_{n,y}$ and $C_{n,y}$ are scene dependent.

[0042] Let be $\min \{I_{n,y}\}$ a pixel being being exposed while $P(t)=0$, and $\max \{I_{n,y}\}$ a pixel being exposed while $P(t)=1$, $\max \{I_{n,y}\} = t_e k_{n,y} + C_{n,y}$. Now, we define a normalized image $\tilde{I}_{n,y}$ as,

$$\tilde{I}_{n,y} = \frac{(n - t_{n,y})P_n + (t_{n,y} + t_e - n)P_{n+1}}{t_e}$$

[0043] A normalized image is completely defined by the time variables and pattern values. In this section we want to estimate the time variables. Let's rewrite Equation 58 as

$$\tilde{I}_{n,y} = -n\Delta P_n \frac{1}{t_e} + (n-1)\Delta P_n \frac{t_f}{t_e} + y\Delta P_n \frac{t_r}{t_e} + \Delta P_n \frac{t_0}{t_e} + P_{n+1}$$

[0044] being t_0 and d unknown. Image pixel values are given by

$$I_n(x, y) = (1 - t_0 - yd)P_n(x, y) + (t_0 + yd)P_{n+1}(x, y).$$

[0045] Same as before, $P_n(x, y)$ and $P_{n+1}(x, y)$ represent the pattern values contributing to camera pixel (x, y) , we define $P_{N+1} = P_1$, $P_0 = P_N$, $I_{N+1} = I_1$, and $I_0 = I_N$, and I will omit pixel (x, y) to simplify the notation. We now minimize the following energy to find the time variables t_0 and d

$$E(t_0, d) = \frac{1}{2} \sum_{n=1}^N \sum_{x,y} ((1 - t_0 - yd)P_n + (t_0 + yd)P_{n+1} - I_n)^2$$

[0046] The partial derivatives are given by

$$\frac{\partial E(t_0, d)}{\partial t_0} = \sum_{n=1}^N \sum_{x,y} (P_{n+1} - P_n)((1 - t_0 - yd)P_n + (t_0 + yd)P_{n+1} - I_n)$$

$$\frac{\partial E(t_0, d)}{\partial d} = \sum_{n=1}^N \sum_{x,y} y(P_{n+1} - P_n)((1 - t_0 - yd)P_n + (t_0 + yd)P_{n+1} - I_n)$$

[0047] We set the gradient equal to the null vector and reorder as

$$\begin{bmatrix} t_0 \\ d \end{bmatrix} = \left(\sum_{n,x,y} (P_{n+1} - P_n)^2 \begin{bmatrix} 1 & y \\ y & y^2 \end{bmatrix} \right)^{-1} \left(\sum_{n,x,y} (P_{n+1} - P_n)(I_n - Pn) \begin{bmatrix} 1 \\ y \end{bmatrix} \right)$$

[0048] We use Equation 29 to compute t_0 and d when we have some known (or estimated) pattern values.

[0049] With known t_0 and d we estimate pattern values minimizing

$$E(P_1, \dots, P_N) = \frac{1}{2} \sum_{n=1}^N ((1 - t_0 - yd)P_n + (t_0 + yd)P_{n+1} - I_n)^2.$$

[0050] Analogous as in Case 1 we obtain that $A\mathbf{p}=\mathbf{b}$ with A as in Equation 12 and α , β , and \mathbf{b} defined as

$$\alpha = (1 - t_0 - yd)^2 + (t_0 + yd)^2, \beta = (1 - t_0 - yd)(t_0 + yd)$$

$$\mathbf{b} = (1 - t_0 - yd)(I_1 I_2 \dots I_N)^T + (t_0 + yd)(I_N I_1 \dots I_{N-1})^T$$

[0051] Pattern values for each pixel are given by $\mathbf{p} = A^{-1} \mathbf{b}$.

[0052] In another preferred embodiment, the method to synthesize the synchronized sequence of image frames from an unsynchronized sequence of image frames, applies to a global shutter image sensor where the image frame rate is higher or equal than the pattern frame rate. FIG. 6 illustrates the timing for this embodiment. We now project M patterns at fixed framerate and we capture N images with a global shutter camera, also at a fixed framerate. We require that $N \geq M$. Capture begins while pattern P_1 is being projected.

We introduce a new variable d which is the camera capture delay from one row to the next. Same as in Case 1, up to two patterns may contribute to each image but here we do not know which ones are because the camera framerate is unknown. The new image equation is

$$I_n = \int_{t_{n-1}}^{t_n} \sum_{m=1}^M f_m(t) P_m dt = \sum_{m=1}^M P_m \int_{t_{n-1}}^{t_n} f_m(t) dt$$

$$f_m(t) = \begin{cases} 1 & \text{if } m-1 \leq t \leq m \\ 0 & \text{otherwise} \end{cases}$$

[0053] Let be $\Delta t = t_{n-1} - t_n$ the time between image frames, let be $\mathbf{p} = (P_1, \dots, P_M)^T$ and $\Phi_n(t_0, \Delta t) = (\Phi(n, 1, t_0, \Delta t), \dots, \Phi(n, M, t_0, \Delta t))^T$ and rewrite Equation 33 as

$$I_n = \frac{1}{\Delta t} \Phi_n(t_0, \Delta t)^T \mathbf{p}$$

[0054] Each function $\Phi(n, m, t_0, \Delta t) = \int_{t_{n-1}}^{t_n} f_m(t) dt$ can be written as

$$\Phi(n, m, t_0, \Delta t) = \max(0, \min(m, t_n) - \max(m-1, t_{n-1}))$$

[0055] Same as before, $P_n(x, y)$ represents a pattern value contributing to camera pixel (x, y) , we define $P_{N+1} = P_1$, $P_0 = P_N$, $I_{N+1} = I_1$, and $I_0 = I_N$, and I will omit pixel (x, y) to simplify the notation.

[0056] We now minimize the following energy to find the time variables t_0 and Δt

$$E(t_0, \Delta t) = \frac{1}{2} \sum_{n=1}^N \sum_{x,y} \left(\frac{1}{\Delta t} \Phi_n(t_0, \Delta t)^T \mathbf{p} - I_n \right)^2$$

[0057] We solve for t_0 and Δt by making $\nabla E(t_0, \Delta t) = 0$

$$\nabla E(t_0, \Delta t) = \sum_{n=1}^N \sum_{x,y} \frac{1}{\Delta t} J \Phi_n(t_0, \Delta t)^T \mathbf{p} \left(\frac{1}{\Delta t} \Phi_n(t_0, \Delta t)^T \mathbf{p} - I_n \right)$$

[0058] Because $J \Phi_n(t_0, \Delta t)$ depends on the unknown value $\mathbf{t} = (t_0, \Delta t)^T$ we solve for them iteratively

$$t^{(i+1)} = A_i(t^{(i)})^{-1} b_i(t^{(i)})$$

$$A_i(t) = \sum_{n=1}^N \sum_{x,y} \frac{1}{\Delta t^2} J \Phi_n(t)^T \mathbf{p} \mathbf{p}^T V_A(n, t)$$

$$b_i(t) = \sum_{n=1}^N \sum_{x,y} \frac{1}{\Delta t} J \Phi_n(t)^T \mathbf{p} \left(I_n - \frac{1}{\Delta t} V_b(n, t)^T \mathbf{p} \right)$$

[0059] Matrix $V_A(n, t)$ and vector $V_b(n, t)$ are defined such as

$$\phi_n(t) = V_A(n, t) \begin{bmatrix} t_0 \\ \Delta t \end{bmatrix} + V_b(n, t)$$

[0060] For completeness we include the following definitions:

$$V_A(n, t) = [v_A(n, 1, t), \dots, v_A(n, M, t)]^T$$

$$V_b(n, t) = [v_b(n, 1, t), \dots, v_b(n, M, t)]^T$$

$$t_{diff} \equiv \min(m, t_n) - \max(m-1, t_{n-1})$$

$$v_A(n, m, t) = \begin{cases} [0, 1]^T & \text{if } m-1 \leq t_{n-1} \leq t_n \leq m \text{ and } t_{diff} > 0 \\ [-1, 1-n]^T & \text{if } m-1 \leq t_{n-1} \leq m \leq t_n \text{ and } t_{diff} > 0 \\ [1, n]^T & \text{if } t_{n-1} \leq m-1 \leq t_n \leq m \text{ and } t_{diff} > 0 \\ [0, 0]^T & \text{if } t_{n-1} \leq m-1 \leq m \leq t_n \text{ and } t_{diff} > 0 \\ [0, 0]^T & \text{otherwise} \end{cases}$$

$$v_b(n, m, t) = \begin{cases} 0 & \text{if } m-1 \leq t_{n-1} \leq t_n \leq m \text{ and } t_{diff} > 0 \\ m & \text{if } m-1 \leq t_{n-1} \leq m \leq t_n \text{ and } t_{diff} > 0 \\ 1-m & \text{if } t_{n-1} \leq m-1 \leq t_n \leq m \text{ and } t_{diff} > 0 \\ 1 & \text{if } t_{n-1} \leq m-1 \leq m \leq t_n \text{ and } t_{diff} > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$J\phi_n(t) = V_A(n, t)$$

[0061] With known t_0 and $\mathbf{66} t$ we estimate pattern values minimizing

$$E(p) = \frac{1}{2} \sum_{n=1}^N (\phi_n(t)^T p - I_n)^2$$

[0062] Analogous as in Case 1 we obtain that $A p = b$ with

$$A = \Phi^T \Phi,$$

$$b = \Phi^T I$$

$$\Phi = \begin{bmatrix} \phi_1(t)^T \\ \vdots \\ \phi_N(t)^T \end{bmatrix}$$

[0063] Pattern values for each pixel are given by $p = A^{-1}b$.

[0064] In another preferred embodiment, the method to synthesize the synchronized sequence of image frames from an unsynchronized sequence of image frames, applies to a rolling shutter image sensor where the image frame rate is higher or equal than the pattern frame rate. FIG. 7 illustrates the timing for this embodiment. Projector framerate is 1, pattern P_m is projected between time $m-1$ and m . A camera frame is read every t_f time. A camera row requires t_r time to be readout from the sensor, thus, a sensor with Y rows needs a time $Y t_r$ to read a complete frame, $t_f \geq Y t_r$. Each camera frame is exposed t_e time, its readout begins immediately after exposure ends, $t_e + t_r \leq t_f$.

[0065] Camera row y in image n begins being exposed at time $t_{n,y}$

$$t_{n,y} = t_0 + (n-1)t_f + y t_r, y: 0 \dots Y-1$$

[0066] and exposition ends at time $t_{n,y} + t_e$.

[0067] In this model a pixel intensity in image n at row y is given by

$$I_{n,y} = \int_{t_{n,y}}^{t_{n,y} + t_e} k_{n,y} P(t) dt + C_{n,y}$$

$$I_{n,y} = \sum_{m=1}^M \max\left(0, \int_{\max(t_{n,y}, m-1)}^{\min(t_{n,y} + t_e, m)} k_{n,y} P_m dt\right) + C_{n,y}$$

[0068] The constants $k_{n,y}$ and $C_{n,y}$ are scene dependent, P_m is either 0 or 1.

[0069] Let $\min\{I_{n,y}\}$ a pixel being exposed while $P(t)=0$, and $\max\{I_{n,y}\}$ a pixel being exposed while $P(t)=1$,

$$\min\{I_{n,y}\} = C_{n,y}$$

$$\max\{I_{n,y}\} = t_e k_{n,y} + C_{n,y}$$

[0070] Now, we define a normalized image $\tilde{I}_{n,y}$ as,

$$\tilde{I}_{n,y} = \frac{I_{n,y} - \min\{I_{n,y}\}}{\max\{I_{n,y}\} - \min\{I_{n,y}\}}$$

$$\tilde{I}_{n,y} = \frac{1}{t_e} \sum_{m=1}^M \max\left(0, \int_{\max(t_{n,y}, m-1)}^{\min(t_{n,y} + t_e, m)} P_m dt\right)$$

[0071] A normalized image is completely defined by the time variables and pattern values. In this section we want to estimate the time variables. Let's rewrite the previous equation as,

$$\tilde{I}_{n,y} = \sum_{m=1}^M \phi(n, m, y) P_m$$

$$\phi(n, m, y) = \frac{1}{t_e} \max(0, \min(t_{n,y} + t_e, m) - \max(t_{n,y}, m-1))$$

$$\text{Let be } h = \frac{1}{t_e} (1, t_f, t_r, t_0)^T,$$

now we write $\phi(n, m, y) = v_{nmy}^T h$ with v_{nmy} defined as

$$v_{nmy}^T = \begin{cases} [t_e, 0, 0, 0] & \text{if } t_{n2} < t_{m2} \wedge t_{n1} > t_{m1} \wedge t_{n2} > t_{n1} \\ [t_e - m + 1, n - 1, y, 1] & \text{if } t_{n2} < t_{m2} \wedge t_{n1} \leq t_{m1} \wedge t_{n2} > t_{n1} \\ [m, 1 - n, -y, -1] & \text{if } t_{n2} \geq t_{m2} \wedge t_{n1} > t_{m1} \wedge t_{m2} > t_{n1} \\ [1, 0, 0, 0] & \text{if } t_{n2} \geq t_{m2} \wedge t_{n1} \leq t_{m1} \wedge t_{m2} > t_{n1} \\ [0, 0, 0, 0] & \text{otherwise} \end{cases}$$

$$t_{n1} = t_{n,y}, t_{n2} = t_{n,y} + t_e, t_{m1} = m - 1, t_{m2} = m$$

Let be $p = (P_1, \dots, P_M)^T$ and let V_{ny} be

$$V_{ny} = \begin{bmatrix} -v_{n1y}^T & - \\ \vdots & \\ -v_{nMy}^T & - \end{bmatrix}$$

[0072] We now minimize the following energy to find the unknown h

$$E(h) = \frac{1}{2} \sum_{n=1}^N \sum_{x,y} (p(x, y)^T V_{n,y} h - \tilde{I}_n(x, y))^2$$

[0073] with the following constraints

$$h > \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \text{ and } \begin{bmatrix} 0 & -1 & Y & 0 \\ 0 & -1 & 1 & 0 \end{bmatrix} h \leq \begin{bmatrix} 0 \\ 0 \\ -1 \end{bmatrix}$$

[0074] or equivalently

$$\frac{t_r}{e} - \frac{t_f}{t_e} \leq -1 \Rightarrow t_r + t_e \leq t_f$$

[0075] Equation $E(h)$ cannot be minimized in closed form because the values matrix $V_{n,y}$ depends on the unknown values. Using an iterative approach the current value $h^{(i)}$ is used to compute $V_{n,y}^{(i)}$ and the next value $h^{(i+1)}$.

[0076] Up to this point we have assumed that the only unknown is h , meaning that pattern values are known for all image pixels. The difficulty lies in knowing which pattern pixel is being observed by each camera pixel. We simplify this issue by making calibration patterns all ‘black or all ‘white’, best seen in FIG. 8. For example, a sequence of four patterns ‘{black, black, white, white}’ will produce images with completely black and completely white pixels, as well as pixels in transition from black to white and vice versa. The all black or white pixels are required to produce normalized images, as shown in FIG. 9, and the pixels in transition constrain the solution of the parameter h in Equation $E(h)$.

[0077] Decoding is done in two steps: 1) the time offset t_0 need to be estimated for this particular sequence; 2) the pattern values are estimated for each camera pixel, as shown in FIG. 10. Value t_0 is estimated using Equation $E(h)$ where the known components of h are fixed, but some pattern values are required to be known, specially we need to know for some pixels whether they are transitioning from ‘black’ to ‘white’ or the opposite. Non-transitioning pixels provided no information in this step. Until now, we have projected a couple of black’s and white’s at the beginning of the sequence to ensure we can normalized all pixels correctly and to simplify t_0 estimation. We will revisit this point in the future for other pattern sequences.

[0078] Similarly as for the time variables, pattern values are estimated by

[0079] minimizing the following energy

$$E(p) = \frac{1}{2} \sum_{n=1}^N (h^T V_{n,y}^T p(x, y) - \tilde{I}_n(x, y))^2, \text{ s.t. } |p_m| \leq 1$$

[0080] The matrix $h^T V_{n,y}^T$ is bi-diagonal for $N=M$ and it is fixed if h is known.

[0081] Therefore, it can be seen that the exemplary embodiments of the method and system provides a unique solution to the problem of using structure lighting for three-dimensional image capture where the camera and projector are unsynchronized.

[0082] It would be appreciated by those skilled in the art that various changes and modifications can be made to the illustrated embodiments without departing from the spirit of the present invention. All such modifications and changes are intended to be within the scope of the present invention except as limited by the scope of the appended claims.

[0083] What is claimed is:

1. A system to capture the surface geometry a three-dimensional object in a scene, comprising:

a pattern projector configured and arranged to project a sequence of image patterns onto the scene at a pattern frame rate;

a camera configured and arranged to capture a sequence of unsynchronized image patterns of the scene at an image capture rate; and

a processor configured and arranged to synthesize a sequence of synchronized image frames from the unsynchronized image patterns of the scene, each of the synchronized image frames corresponding to one image pattern of the sequence of image patterns.

2. The system of claim 1, wherein the sequence of image patterns comprises binary patterns.

3. The system of claim 1, wherein the number of image patterns in the sequence of image patterns is less than or equal to the number of unsynchronized image frames in the sequence of unsynchronized image frames.

4. The system of claim 1, wherein the camera has a rolling shutter operation.

5. The system of claim 1, wherein the camera has a global shutter operation.

6. The system of claim 1, wherein the image capture rate of the camera is equal to the pattern frame rate of the projector.

7. The system of claim 1, wherein the image capture rate of the camera is greater than the pattern frame rate of the projector.

8. A method of capturing the surface geometry of a three-dimensional object in a scene, comprising:

projecting a sequence of image pattern into the scene at a pattern frame rate;

capturing a sequence of unsynchronized image patterns of the scene at an image capture rate; and

synthesizing a sequence of synchronized image frames from the from the unsynchronized image patterns of the scene, each of the synchronized image frames corresponding to one image pattern of the sequence of image patterns.

9. The method of claim 8, wherein the step of projecting a sequence of image patterns comprises projecting a sequence of binary patterns.

10. The method of claim 8, wherein the number of image patterns projected in the sequence of image patterns is less than or equal to the number of unsynchronized image frames in the sequence of unsynchronized image frames.

11. The method of claim 8, further comprising selecting a pattern frame rate.

12. The method of claim 8, further comprising selecting an image capture rate.

13. The method of claim **8**, wherein the image capture rate is equal to the pattern frame rate.

14. The method of claim **8**, wherein the image capture rate is greater than the pattern frame rate.

15. The method of claim **8**, further comprising decoding the sequence of synchronized image frames.

16. The method of claim **15**, further comprising applying three-dimensional triangulation to the sequence of synchronized image frames.

17. The method of claim **16**, further comprising applying geometric processing to the sequence of synchronized image frames.

18. A method to synthesize a sequence of synchronized image frames synchronized from a sequence of unsynchronized image frames; each of the synchronized image frames corresponding to one of a sequence of image patterns; the synchronized image frames and unsynchronized image frames being image frames of the same width and height; the image frames comprising a plurality of common pixels; the method comprising the steps of:

- a) partitioning the plurality of common pixels into a pixel partition; the pixel partition comprising a plurality of pixel sets; the pixel sets being disjoint; each of the plurality of common pixels being a member of one of the pixel sets;
- b) selecting a pixel set;
- c) building a pixel set measurement matrix;

d) estimating a pixel set projection matrix; the pixel set projection matrix projecting a measurement space vector onto the column space of the pixel set measurement matrix;

e) estimating a pixel set system matrix; the system matrix being parameterized by a set of system matrix parameters;

f) estimating a pixel set synchronized matrix as a function of the pixel set measurement matrix and the pixel set system matrix;

g) repeating steps b) to f) until all the pixels sets have been selected; and

h) constructing a sequence of synchronized image frames from the pixel set synchronized matrices.

19. A method as in claim **18**, where the pixel partition comprises a single pixel set, and the single pixel set contains all the pixels of the image frames.

20. A method as in claim **18**, where the number of pixel sets in the pixel partition is equal to the height of the image frames, and each row of the image frames is a pixel set.

21. A method as in claim **18**, where the pixel set measurement matrix is modeled as the product of the system matrix times the pixel set synchronized matrix, and the step of estimating the pixel set synchronized matrices reduces to the solution of a linear least-squares problem.

22. A method as in claim **18**, where the system matrix is parameterized by an image frame period parameter, an integration time parameter, and a first pattern delay parameter.

* * * * *