(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
31 December 2014 (31.12.2014)

WIPO | PCT

(10) International Publication Number
WO 2014/205643 A1

(51) International Patent Classification:
H04N 13/00 (2006.01) *G06T 7/00* (2006.01)

(21) International Application Number:
PCT/CN2013/077844

(22) International Filing Date:
25 June 2013 (25.06.2013)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant *(for all designated States except US)*: **THOMSON LICENSING** [FR/FR]; 1-5 rue Jeanne d'Arc, F-92130 Issy-les-Moulineaux (FR).

(72) Inventors; and

(71) Applicants *(for US only)*: **SONG, Jianping** [CN/CN]; 8th Floor, Building A, Technology Fortune Center, No. 8 Xueqing Road, Haidian District, Beijing 100192 (CN). **WANG, Shilin** [CA/CN]; 8th Floor, Building A, Technology Fortune Center, No. 8 Xueqing Road, Haidian District, Beijing 100192 (CN). **DU, Lin** [CN/CN]; 8th Floor, Building A, Technology Fortune Center, No. 8 Xueqing Road, Haidian District, Beijing 100192 (CN). **MA, Xiaojun** [CN/CN]; 8th Floor, Building A, Technology Fortune Center, No. 8 Xueqing Road, Haidian District, Beijing 100192 (CN).

(74) Agent: **LIU, SHEN & ASSOCIATES**; A0601, Huibin Building, No. 8 Beichen Dong Street, Chaoyang District, Beijing 100101 (CN).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, GH,

[Continued on next page]

(54) Title: METHOD AND SYSTEM CAPABLE OF ALIGNMENT OF VIDEO FRAME SEQUENCES

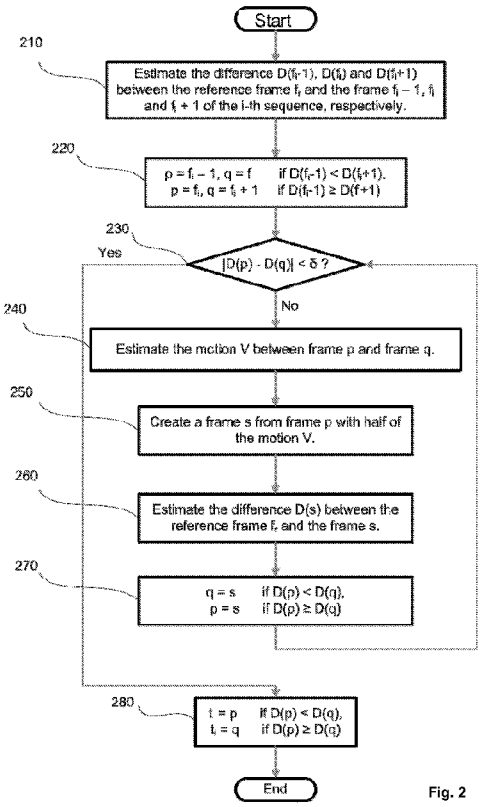

Fig. 2

(57) Abstract: The present invention is related to a method for alignment of video frame sequences. The method comprises the steps of: determining an initial aligned frame in a video frame sequence with respect to a reference frame in a reference video frame sequence; and creating a virtual frame in the video frame sequence which virtual frame is more temporally aligned to the reference frame than the initial aligned frame. The created virtual frame is selected as the frame in the video frame sequence corresponding to the reference frame in the reference video frame sequence.

WO 2014/205643 A1

GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:

— *with international search report (Art. 21(3))*

METHOD AND SYSTEM CAPABLE OF ALIGNMENT OF VIDEO FRAME SEQUENCES

TECHNICAL FIELD

The present invention relates to a method and system capable of alignment of video frame sequences.

BACKGROUND ART

Multiple cameras can be utilized in a number of applications such as multi-view TV, free viewpoint TV, 3D teleconference and 3D surveillance. In a 3D application, an accurate 3D scene can be presented only when the image of the left view and the image of the right view are of the same scene captured at the same time. However, in an implementation of actual 3D applications, temporal latency and spatial shift may be introduced between video frame sequences. Typically, a temporal misalignment between video frame sequences occurs when the input sequences have different frame rates (e.g., NTSC and PAL), or when there is a time shift between the sequences (e.g. when the cameras are not activated simultaneously). On the other hand, a spatial misalignment results from the different positions, orientations and internal calibration parameters of all the cameras. Therefore, it would be necessary to establish both the temporal synchronization (temporal alignment) and the spatial alignment between the video frame sequences. This establishment would be required in applications such as tele-immersion, video-based surveillance, panoramic video mosaicing and video metrology.

While a hardware synchronization method can embed timestamps into each video frame sequence on-the-fly and requires no post-processing, it requires a specialized hardware as well as setting up a camera network in advance. On the other hand,

a computer vision-based software synchronization algorithm can be used to perform post-processing on video frame sequences recorded by cameras that are not networked, such as common consumer hand-held video cameras or cameras embedded in mobile phones, or to synchronize historical videos for which hardware synchronization cannot be applicable.

The alignment may include shifting and correcting spatial and temporal detection of the video sequences. Many alignment methods exist, one of which is a spatio-temporal alignment method. The spatio-temporal alignment method can be divided into two main classes: a feature-based method and a direct method. The feature-based method uses detected features as the main input for alignment (e.g., two-frame feature correspondences or multi-frame feature trajectories). From the feature correspondences or trajectory, an algebraic or geometric error is computed and used as a measure of synchronization. On the other hand, the direct method relies on colors, intensities and intensity gradients to determine the spatio-temporal alignment of overlapping video frame sequences. As a result, the direct method tends to align sequences more accurately when their intensities are similar, while the feature-based method is appropriate when scene appearance varies greatly from sequence to sequence (e.g., due to wide baselines, different magnification, or cameras with distinct spectral sensitivities).

Existing alignment methods merely implement inter-frame alignment. That is, for each frame in one video sequence, a corresponding frame in the other video sequence is found and transformed so that the two frames of the frame in the one video sequence and the corresponding frame in the other video sequence are aligned both in time and space.

The frame rate of the video frame sequence is typically 25 or 30 frames per second. This means there may be a time shift of

0.5 frames up to 20 milliseconds (= 1/25 x 1000 / 2) between two aligned frames. For some applications, such a time shift may cause problems. For example, in a 3D reconstruction application, the depth map should be built based on left and right views. A time shift of 0.5 frames up to 20 milliseconds between the left and right views may impact the preciseness of the depth map observably.

A method for video synchronization is shown in the technical paper: Xiaochun Cao, Lin Wu, Jiangjian Xiao, H. Foroosh, Jigui Zhu and Xiaohong Li, "Video synchronization and its application to object transfer", Image and Vision Computing, 2010.

A method for video frame alignment is shown in the technical paper: S. Kuthirummal, C.V. Jawahar and P.J. Narayanan, "Video frame alignment in multiple views", Proceedings 2002 International Conference on Image Processing, 2002.

Also, a method for spatio-temporal alignment of sequences is shown in the technical paper: Y. Capsi and M. Irani, "Spatio-temporal alignment of sequences", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE 24 (11), page 1409-1424, NOVEMBER 2002.

SUMMARY OF THE INVENTION

According to one aspect of the present invention, a method for alignment of video frame sequences may comprise the steps of: determining an initial aligned frame in a video frame sequence with respect to a reference frame in a reference video frame sequence; and creating a virtual frame in the video frame sequence which virtual frame is more temporally aligned to the reference frame than the initial aligned frame. The created virtual frame may be selected as the frame in the video frame sequence corresponding to the reference frame in

the reference video frame sequence.

According to another aspect of the present invention, a system for alignment of video frame sequences may comprise a processor which is configured to perform the processes of: determining an initial aligned frame in a video frame sequence with respect to a reference frame in a reference video frame sequence; and creating a virtual frame in the video frame sequence which virtual frame is more temporally aligned to the reference frame than the initial aligned frame. The created virtual frame may be selected as the frame in the video frame sequence corresponding to the reference frame in the reference video frame sequence.

The object and advantages of the invention will be realized and attained by means of the elements and combinations particularly pointed out in the claims.

It is to be understood that both the foregoing general description and the following detailed description are exemplary and explanatory and are not restrictive of the invention, as claimed.

BRIEF DESCRIPTION OF THE DRAWINGS

These and other aspects, features and advantages of the present invention will become apparent from the following description in connection with the accompanying drawings in which:

Fig. 1A illustrates an example of an array of video cameras shooting the same scene;

Fig. 1B illustrates the reference frame f_r in the reference sequence as well as the nearest initial aligned frame f_i with respect to the reference frame f_r and the created virtual

frame t_i in the i-th sequence;

FIG. 2 is a flow chart illustrating the processes for alignment of video frame sequences according to an embodiment of the present invention; and

Fig. 3 is a block diagram schematically illustrates a system 300 for performing the processes for the alignment of video frame sequences according to an embodiment of the present invention.

DETAILED DESCRIPTION

In the following description, various aspects of an embodiment of the present invention will be described. For the purpose of explanation, specific configurations and details are set forth in order to provide a thorough understanding. However, it will also be apparent to one skilled in the art that the present invention may be implemented without the specific details present herein.

Fig. 1A illustrates an example of an array of video cameras shooting the same scene.

In the example shown in Fig. 1A, a single dynamic scene 105 is shot simultaneously by N respective cameras 110a, 110b, 110c, ... 110n arranged at distinct viewpoints. It is assumed that each camera captures frames with a constant, but unknown frame rate. It is also assumed that the cameras are unsynchronized, i.e., they began capturing video frames at different timings with possibly distinct frame rate. In order to temporally align the resulting video frame sequences, the correspondence relationship between frame numbers in one "reference" sequence and frame numbers in all other sequences should be determined. This correspondence can be expressed as a set of linear equations,

$$t_i = \alpha_i f_r + \beta_i \quad \text{Eq. (1)}$$

where f_r denotes the frame numbers of the reference sequence, and t_i denote the aligned frame numbers of the i-th sequence. α_i and β_i are unknown constants representing temporal dilation and temporal shift, respectively, between the i-th sequence and the reference sequence.

In general, α_i and β_i are not necessarily integers. However, f_r must be an integer. As a result, the computed frame number t_i is not necessarily an integer. If t_i is not an integer, it implies that the most temporally aligned frame of the i-th sequence, with respect to reference frame f_r, is not captured by camera i. Therefore, frame t_i may be a virtual frame that is not of the sequence captured by camera i. However, existing alignment techniques just handle captured frames. Therefore, existing alignment techniques may find that frame f_i is the nearest frame to the reference frame f_r, where f_i usually is the rounding of t_i to the nearest integer. Therefore,

$$f_i \le t_i \le f_i + 1, \quad \text{if } f_i \le t_i \quad \text{Eq. (2)}$$

or

$$f_i - 1 \le t_i \le f_i, \quad \text{if } f_i \ge t_i \quad \text{Eq. (3)}$$

Fig. 1B illustrates the reference frame f_r in the reference sequence as well as the nearest initial aligned frame f_i with respect to the reference frame f_r and the created virtual frame t_i in the i-th sequence, which case is represented by the above described equation (3).

In an embodiment of the present invention, an existing video alignment technique or a combination of existing video alignment techniques is performed to determine the initial frame f_i.

A feature-based alignment approach is one of such video alignment techniques. This approach comprises: to first extract distinctive features from each frame, to match these features in order to establish a global correspondence, and then to estimate the geometric transformation between the feature images. This kind of approach has been used since the early days of stereo matching and has more recently gained popularity for image stitching applications.

Another video alignment technique is known as a direct alignment. This technique uses pixel intensities in a video frame for synchronization and is suitable for videos containing significant lighting changes, e.g., fireworks, or flickering fire. For example, a direct alignment algorithm may synchronize two sequences in a coarse-to-fine manner. Firstly, a Gaussian sequence pyramid is computed, which is the video sequence equivalent of a Gaussian image pyramid. At each level of the pyramid, an iterative algorithm minimizes the sum-of-squared differences in pixel intensities between sequences according to the current estimate of the spatio-temporal model.

Once the initial aligned frame f_i is found, an iterative procedure is performed to refine the alignment and to look for the most aligned virtual frame t_i.

Fig. 2 is flow chart illustrating the processes for alignment of video frame sequences according to an embodiment of the present invention.

At step 210, the differences $D(f_i - 1)$, $D(f_i)$ and $D(f_i + 1)$ between the reference frame f_r and the initial aligned frame $f_i - 1$, f_i and $f_i + 1$ of the i-th sequence are computed, respectively. Note that frame $f_i - 1$ and $f_i + 1$ are the previous frame and the next frame of frame f_i, respectively.

The difference is a numerical value that indicates the degree of the difference between the two frames, f_r and $f_i - 1$, f_r and f_i and f_r and $f_i + 1$. According to an embodiment of the present invention, any number of difference estimation techniques, or combination of techniques, may be used to determine the difference values.

One of such difference estimation techniques is known as the pixel-wise difference magnitude ("PDM"). The PDM calculates the sum of the difference of all pixel values of two corresponding frames, such as the frame f_i and the reference frame f_r. The pixel difference is typically squared, or the absolute value is taken, to result in a positive magnitude before summation.

Another difference estimation technique is known as the normalized pixel difference ("NPD"). The NPD determines difference by normalizing or scaling the pixels from both source frames to correct for camera intensity or color differences. Then the difference magnitude may be calculated using another difference estimation technique, such as the PDM.

Another technique is known as the frequency-domain comparisons ("FDC"). Using this technique, a spatial transform such as a two-dimensional Fourier transform or wavelet transform may be used to parameterize image regions into transform coefficients. The magnitude difference can be computed between the transform coefficients, which can be selected or weighted to favor or discard particular spatial frequencies.

At step 220, the frame with less difference value is selected from frame $f_i - 1$ and $f_i + 1$. The selected frame as well as frame f_i is denoted as frame p and frame q, respectively. Note that frame p is temporally prior to frame q. Therefore,

if the difference value $D(f_i - 1)$ is less than $D(f_i + 1)$, frame p denotes frame f_i-1 and frame q denotes frame f_i. On the other hand, if the difference value $D(f_i - 1)$ is greater than $D(f_i + 1)$, frame p denotes frame f_i and frame q denotes frame f_i +1.

At step 230, the terminative condition of the iterative process is checked. If the difference between difference values D(p) and D(q) is less than a predefined threshold value δ ("Yes"), it means that frame p and frame q are similar enough. Therefore, the process to refine the alignment can terminate and then the process is passed to step 280 to select the most aligned virtual frame t_i. If the difference between difference values D(p) and D(q) is greater than the predefined threshold value δ ("No"), the process will continue and try to find a more aligned frame. The predefined threshold value δ may be changed by a user by a user interface of the system described below.

At step 240, the motion between frame p and frame q is estimated. Motion estimation is the process of determining motion vectors that describe the transformation from one 2D image to another; usually from adjacent frames in a video sequence. Motion estimation is a key part of video compression as used by MPEG 1, 2 and 4 as well as many other video codecs. It should be noted that the motion needn't be estimated at each iteration. The motion estimated at last iteration can be used to quickly create the motion between frame p and frame q.

At step 250, a motion compensation technique is used to create a frame s from frame p with half of the estimated motion. The motion compensation is a key part of video compression as used by MPEG 1, 2 and 4 as well as many other video codecs.

At step 260, the difference value D(s) between the reference frame f_r and the synthesized frame s is estimated. Note that frame s is synthesized from frame p and the estimation motion, and the difference value D(p) between the reference frame f_r and frame p is already computed. Therefore, the difference value D(s) can be computed from frame p and the estimation, which may be more efficient than to compute D(s) directly from frame s.

At step 270, the denotation of frame number p and q is updated so that the one of frame p and q with greater difference value is discarded. The remainder and the synthesized frame s are denoted as new frame p and q. Still new frame p is temporally prior to new frame q. Therefore, if D(p) is less than D(q), frame q is discarded and new frame q is the synthesized frame s. Otherwise frame p is discarded and new frame p is the synthesized frame s.

Once the denotation of frame number p and q is updated, the process is passed again to step 230 to check the difference between the new frame p and q. If the difference between D(p) and D(q) is less than the predefined threshold value δ ("Yes"), it means that new frame p and new frame q are similar enough. Therefore, the process to refine the alignment can terminate. The process is passed to step 280 to select the most aligned frame t_i. If the difference between D(p) and D(q) is greater than the predefined threshold value δ ("No"), the process will continue and try to find a more aligned frame in another iterative procedure.

Finally at step 280, the frame with less difference value is selected from frame p and q. According to an embodiment of present invention, this frame is the most temporally aligned frame of the i-th video sequence with respect to the reference frame f_r of the reference sequence. This frame is denoted as t_i. This created virtual frame t_i is selected as

the frame in the i-th video frame sequence corresponding to the reference frame f_r in the reference video frame sequence. In the presentation of the i-th video frame sequence, the virtual frame t_i may be presented while the initial aligned frame f_i may be omitted for the presentation.

In the same manner, the most temporally aligned frame of other video sequences with respect to the reference frame f_r of the reference sequence can be found.

In a variant of the embodiment, the alignment process described above may be applied to every frame of the reference video sequence. Therefore, for each frame of the reference sequence, an aligned frame is calculated for each other sequences. As a result, i - 1 new sequences are calculated that are aligned with the reference sequence frame by frame.

In still another variant of the embodiment, instead of computing the aligned frame for each frame of the reference sequence, only a portion of frames of the reference sequence is selected and their corresponding aligned frames are computed. In fact, it can be observed in Eq. (1) that only α_i and β_i are unknown constants. Therefore, if two aligned frames t_i' and t_i'' are already computed with respect to two reference frame f_r' and f_r'', then

$$t_i' = \alpha_i f_r' + \beta_i \quad \text{Eq. (4)}$$

and

$$t_i'' = \alpha_i f_r'' + \beta_i \quad \text{Eq. (5)}$$

According to Eq. (4) and (5), constants α_i and β_i can be calculated. Then for any reference frame f_r, the time of frame t_i can be calculated according to Eq. (1). Therefore for any reference frame f_r, the virtual aligned frame t_i can be created once constants α_i and β_i become known.

In still another variation of the embodiment, instead of using the same sequence as the reference sequence, an already aligned sequence can be used as the reference sequence when aligning another sequence. For example, the system may use sequence i as the reference sequence to align sequence j, then the system use the aligned sequence j as the reference to align sequence k.

Fig. 3 is a block diagram schematically illustrates a system 300 for performing the processes for alignment of video frame sequences according to an embodiment of the present invention as described above in connection with the flow chart shown in Fig. 2.

The system 300 illustrated in Fig. 3 may include a CPU (Central Processing Unit) 310, a storage unit 320, a user interface module 330, and an interface (I/F) module 340 connected via a bus 350. A memory (not shown) such as RAM (Random Access Memory) may be also connected to the CPU 310 via a direct connection or via the bus 350. The connection between the CPU 310 and other parts of the system 300 may be a direct connection, and the connection is not limited to that using the bus 350.

The CPU 310 is an example of a processing unit that controls the processes as described above in connection with the flow chart shown in Fig. 2. The CPU 310 is configured to perform the process of: "video synchronization" for determining an initial frame as discussed above in connection with Figs. 1A and 1B and "frame alignment" of video frame sequences including the steps 210-280 as described above in connection with the flow chart shown Fig. 2.

The storage unit 320 may store one or more programs to be executed by the CPU 310, and various data including the

reference video frame sequence, the other video frame sequences and intermediate data obtained during computations performed by the CPU 310. The storage unit 320 may be formed by a semiconductor memory device, a storage unit that uses a magnetic recording medium, an optical storage unit that uses an optical recording medium, a magneto-optical storage unit that uses a magneto-optic recording medium, or any combination of such devices or units.

Examples of the user interface module 330 may include a keyboard and the like, to be operated by the user when inputting data, instructions, and the like to the system 300. The user can change the threshold value δ with the user interface module 330 as mentioned above in connection with the step 230.

The I/F module 340 may provide an interface between the system 300 and an external apparatus (not illustrated) via a cable interface, a wireless interface, or a combination of cable and wireless interfaces. The I/F 340 module may be connected to a network, such as the Internet, and the system 300 may receive data (for example, the reference video frame sequence and the other video frame sequences as source sequences), instructions, programs, and the like via the I/F module 340. Also, the system 300 may output data such as aligned sequences via the I/F module 340.

A program may cause a computer or a processing unit, such as the CPU 310, to execute the video sequence alignment process. Such a program may be stored in any suitable non-volatile computer-readable storage medium, including a semiconductor memory device, a magnetic recording medium, an optical recording medium, and a magneto-optic recording medium.

All examples and conditional language recited herein are intended for pedagogical purposes to aid the reader in

understanding the invention and the concepts contributed by the inventor to furthering the art, and are to be construed as being without limitation to such specifically recited examples and conditions, nor does the organization of such examples in the specification relate to a showing of the superiority and inferiority of the invention. Although the embodiments of the present invention have been described in detail, it should be understood that the various changes, substitutions, and alterations could be made hereto without departing from the spirit and scope of the invention.

CLAIMS

1. A method for alignment of video frame sequences, comprising the steps of:

determining an initial aligned frame in a video frame sequence with respect to a reference frame in a reference video frame sequence; and

creating a virtual frame in the video frame sequence which virtual frame is more temporally aligned to the reference frame than the initial aligned frame,

wherein the created virtual frame is selected as the frame in the video frame sequence corresponding to the reference frame in the reference video frame sequence.

2. The method as claimed in claim 1, wherein the creating step is repeated to create the virtual frame so that a difference value between the reference frame and a newly created virtual frame becomes smaller than a difference value between the reference frame and a previously created virtual frame until the difference value between the reference frame and the newly created virtual frame becomes smaller than a predetermined value.

3. The method as claimed in claim 2, wherein the virtual frame is created so that a motion between the newly created virtual frame and the previously created frame becomes a half of a motion between the previously created frame and another previously created frame.

4. The method as claimed in claim 1, wherein the determining step and the creating step are applied for every reference frame in the reference sequence to create virtual frames in the video frame sequence corresponding to the reference frames, respectively.

5. A system for alignment of video frame sequences

comprising a processor which is configured to perform the processes of:

determining an initial aligned frame in a video frame sequence with respect to a reference frame in a reference video frame sequence; and

creating a virtual frame in the video frame sequence which virtual frame is more temporally aligned to the reference frame than the initial aligned frame,

wherein the created virtual frame is selected as the frame in the video frame sequence corresponding to the reference frame in the reference video frame sequence.

6. The system as claimed in claim 5, wherein the creating process is repeated to create the virtual frame so that a difference value between the reference frame and a newly created virtual frame becomes smaller than a difference value between the reference frame and a previously created virtual frame until the difference value between the reference frame and the newly created virtual frame becomes smaller than a predetermined value.

7. The system as claimed in claim 6, wherein the virtual frame is created so that a motion between the newly created virtual frame and the previously created frame becomes a half of a motion between the previously created frame and another previously created frame.

8. The system as claimed in claim 5, wherein the determining process and the creating process are applied for every reference frame in the reference sequence to create virtual frames in the video frame sequence corresponding to the reference frames, respectively.

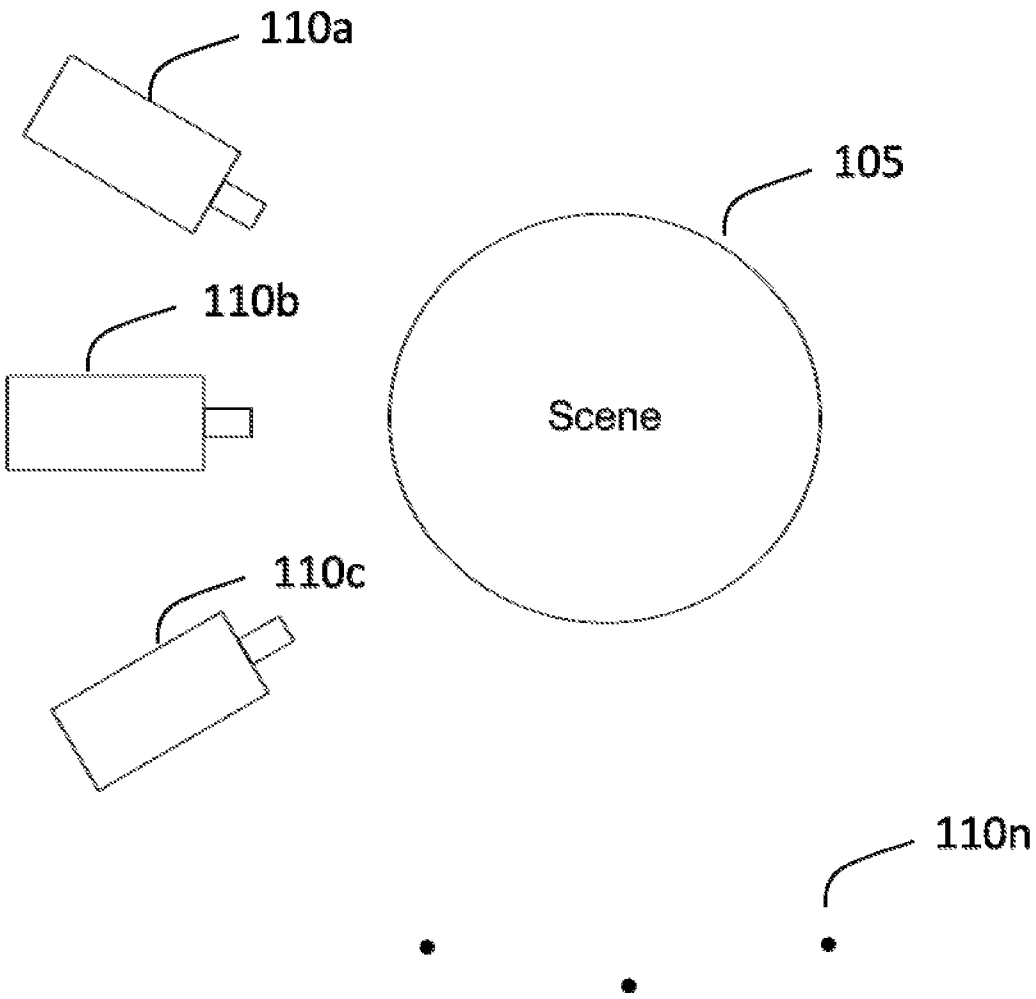

Fig. 1A

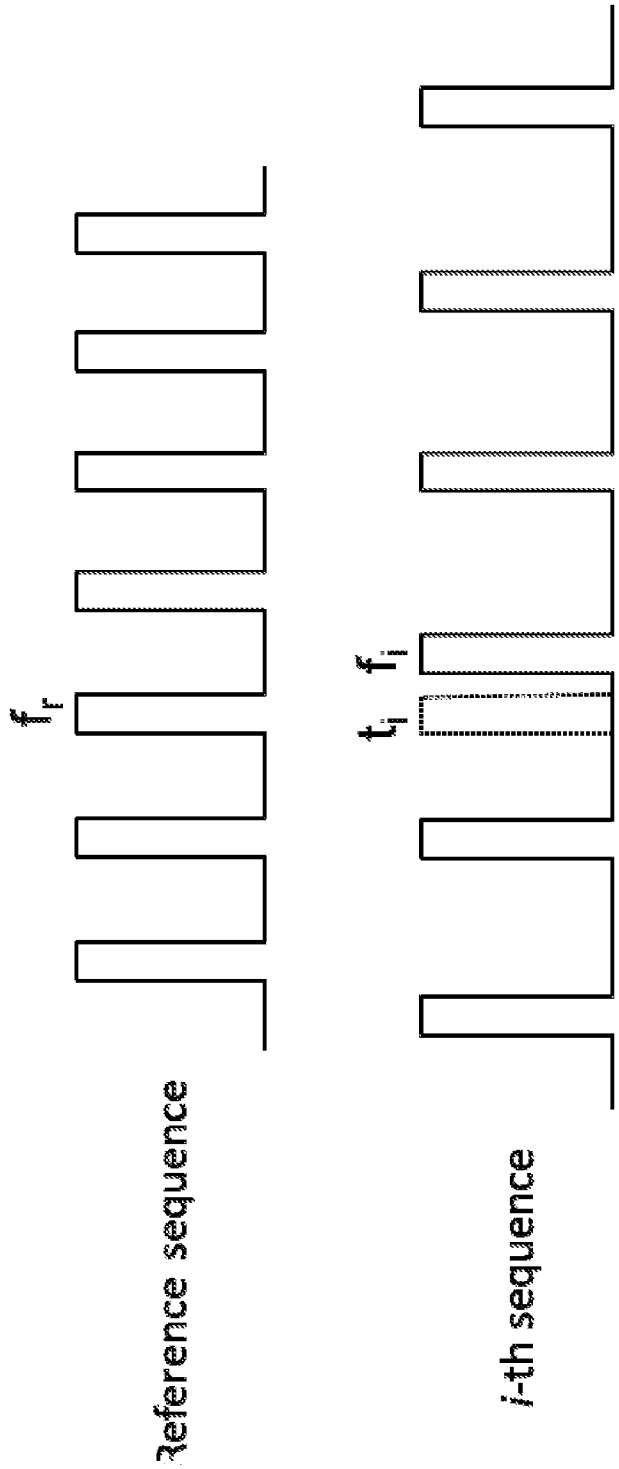

Fig. 1B

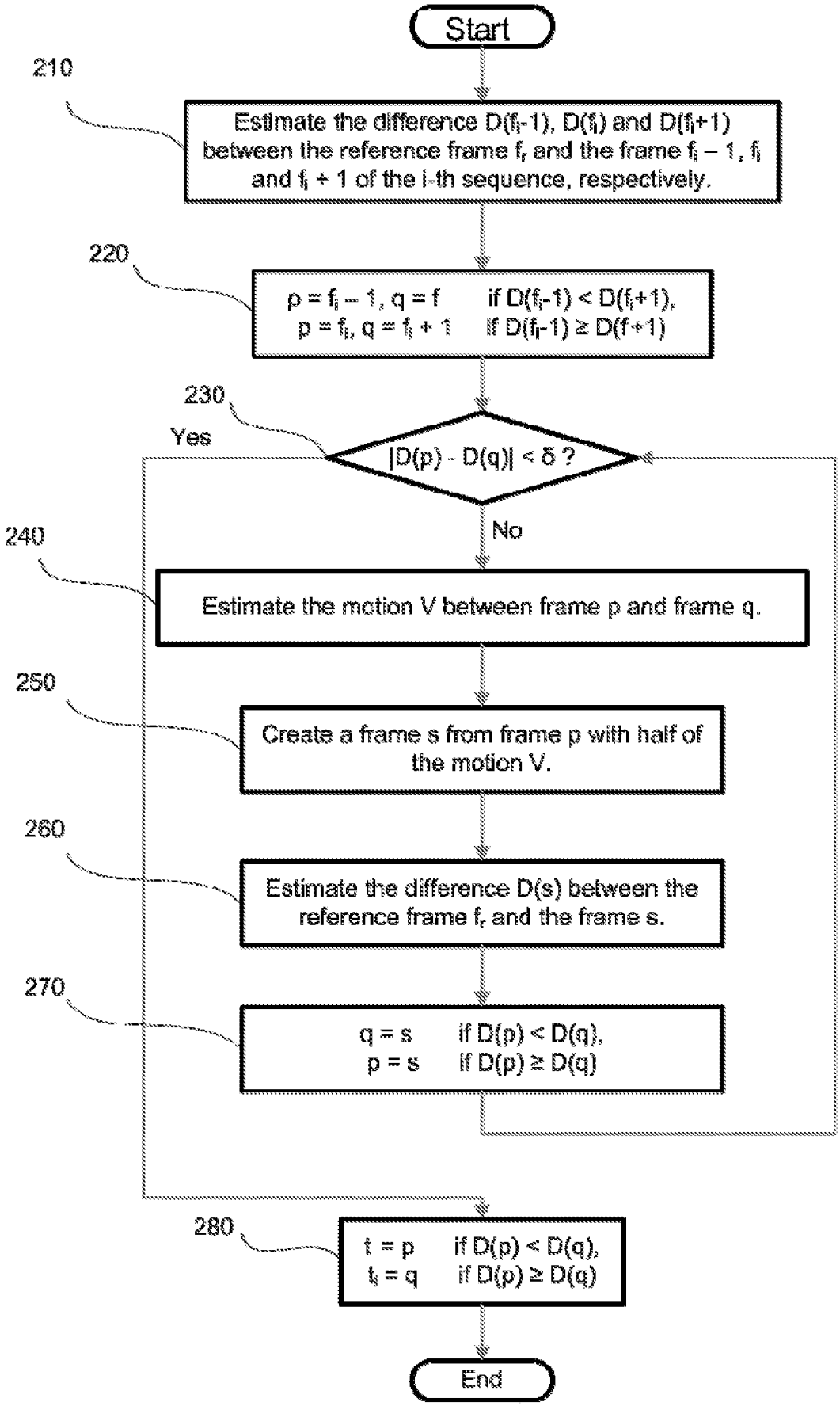
Start
210
Estimate the difference D(f_i-1), D(f_i) and D(f_i+1) between the reference frame f_r and the frame f_i – 1, f_i and f_i + 1 of the i-th sequence, respectively.
220
p = f_i – 1, q = f if D(f_i-1) < D(f_i+1),
p = f_i, q = f_i + 1 if D(f_i-1) ≥ D(f+1)
230
Yes
|D(p) - D(q)| < δ ?
No
240
Estimate the motion V between frame p and frame q.
250
Create a frame s from frame p with half of the motion V.
260
Estimate the difference D(s) between the reference frame f_r and the frame s.
270
q = s if D(p) < D(q),
p = s if D(p) ≥ D(q)
280
t = p if D(p) < D(q),
t_i = q if D(p) ≥ D(q)
End

Fig. 2

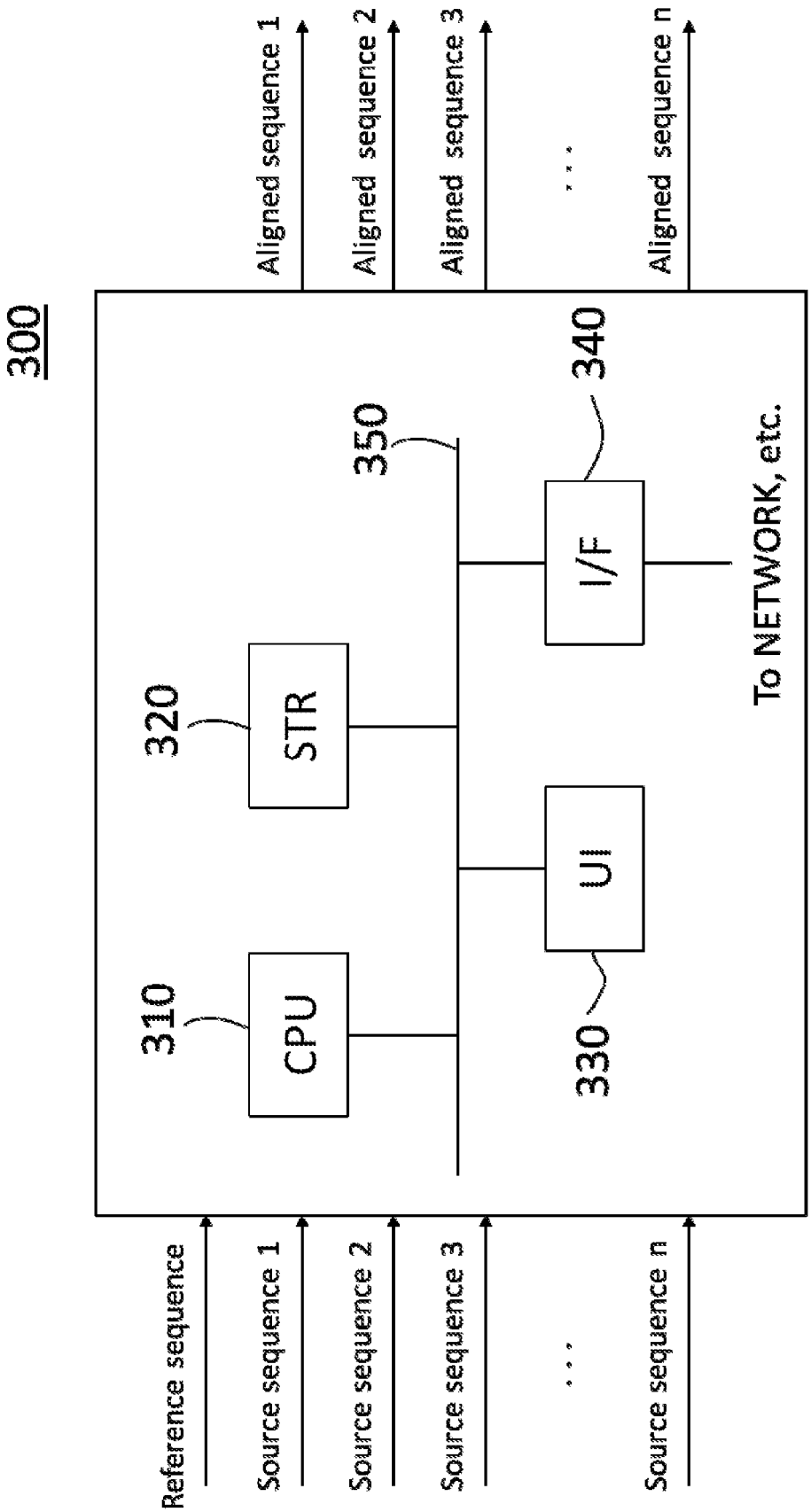
300
Aligned sequence 1
Aligned sequence 2
Aligned sequence 3
. . .
Aligned sequence n
350
340
I/F
To NETWORK, etc.
320
STR
310
CPU
UI
330
Reference sequence
Source sequence 1
Source sequence 2
Source sequence 3
. . .
Source sequence n

Fig. 3

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2013/077844

A. CLASSIFICATION OF SUBJECT MATTER

See the extra sheet

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC: H04N, G06T

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

WPI, EPODOC, CNPAT, CNKI, IEEE: video frame sequence stream, align+ synchroniz+, temporal+ time shift offset, virtual interpolation new, intinal, repeat+ iterative, reference difference

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	US 2011043691 A1 (GUITTENY, Vincent et al.) 24 February 2011 (24.02.2011) description, paragraph [0038] and figures 1-8	1, 4, 5, 8
Y	WO 2012101542 A1 (KONINKLIJKE PHILIPS ELECTRONICS N.V.) 02 August 2012 (02.08.2012) description, page 4, lines 9-26	1, 4, 5, 8
A	US 2012087571 A1 (LEE, Gwang Soon et al.) 12 April 2012 (12.04.2012) the whole document	1-8

☐ Further documents are listed in the continuation of Box C. ☒ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier application or patent but published on or after the international filing date

"L" document which may throw doubts on priority claim (S) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&"document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
27 February 2014 (27.02.2014)	**03 Apr. 2014 (03.04.2014)**
Name and mailing address of the ISA/CN The State Intellectual Property Office, the P.R.China 6 Xitucheng Rd., Jimen Bridge, Haidian District, Beijing, China 100088 Facsimile No. 86-10-62019451	Authorized officer ZHANG, Liguo Telephone No. (86-10)61648239

Form PCT/ISA /210 (second sheet) (July 2009)

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/CN2013/077844

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
US 2011043691 A1	24.02.2011	FR 2922074 A1	10.04.2009
		WO 2009043923 A1	09.04.2009
		FR 2922074 B1	26.02.2010
		CA 2701698 A1	09.04.2009
		EP 2208355 A1	21.07.2010
		EP 2208355 B1	18.05.2011
		AT 510413 T	15.06.2011
WO 2012101542 A1	02.08.2012	US 2013293783 A1	07.11.2013
		CN 103404122 A	20.11.2013
		EP 2668771 A1	04.12.2013
US 2012087571 A1	12.04.2012	KR 20120036724 A	18.04.2012

INTERNATIONAL SEARCH REPORT

International application No.
PCT/CN2013/077844

A. CLASSIFICATION OF SUBJECT MATTER

H04N 13/00 (2006.01) i

G06T 7/00 (2006.01) i