



- (51) **International Patent Classification:**  
*H04N 5/46* (2006.01)
- (21) **International Application Number:**  
PCT/CN2014/080979
- (22) **International Filing Date:**  
27 June 2014 (27.06.2014)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (71) **Applicant:** THOMSON LICENSING [FR/FR]; 1-5 rue Jeanne d' Arc, F-92130 Issy-Les-Moulineaux (FR).
- (72) **Inventors:** WEL, Zhengdong; 8th Floor, Building A, Technology Fortune Center, No. 8 Xueqing Road, Haidian District, Beijing 100192 (CN). YANG, Qinglei; 8th Floor, Building A, Technology Fortune Center, No. 8 Xueqing Road, Haidian District, Beijing 100192 (CN). YU, Qianlong; 8th Floor, Building A, Technology Fortune Center, No. 8 Xueqing Road, Haidian District, Beijing 100192 (CN).
- (74) **Agent:** LIU, SHEN & ASSOCIATES; 10th Floor, Building 1, 10 Caihefang Road, Haidian District, Beijing 100080 (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, IR, 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, SA, 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, 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))

(54) **Title:** METHOD AND DEVICE FOR DISPLAYING A VIDEO SEQUENCE

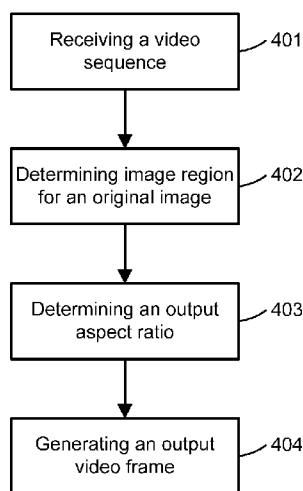


Figure 4

(57) **Abstract:** It is provided a method for displaying a video sequence, comprising the steps of receiving (401) a video frame of the video sequence, wherein the video frame consists of an image and masking; determining (402) an image region of the image within the video frame of the video sequence; determining (403) an output aspect ratio; and generating (404) an output video frame based on the video frame, the image region, the output aspect ratio and a predefined value indicating a filling manner for filling the image into the output video frame.



## METHOD AND DEVICE FOR DISPLAYING A VIDEO SEQUENCE

### TECHNICAL FIELD

The present disclosure relates to image processing, and more particularly  
5 relates to a method and a device for displaying a video sequence.

### BACKGROUND

In technical field of image processing, an image has an aspect ratio. The  
aspect ratio describes the proportional relationship between the image's width  
10 and the image's height, i.e. it is the ratio of width to height.

The most common aspect ratios used today in the presentation of films in  
cinemas are 1.85:1 and 2.39:1. Two common videographic aspect ratios are 4:3  
(1.33:1), which is the universal video format of the 20th century, and 16:9  
(1.77:1), which is universal for high-definition television and European digital  
15 television. Other cinema and video aspect ratios exist, but are used infrequently.  
In still camera photography, the most common aspect ratios are 4:3, 3:2, and  
more recently being found in consumer cameras 16:9. Other aspect ratios, such  
as 5:3, 5:4, and 1:1 (square format), are used in photography as well,  
particularly in medium format and large format.

20

### SUMMARY

According to an aspect of the present disclosure, it is provided a method for  
displaying a video sequence, comprising the steps of receiving (401) a video  
frame of the video sequence, wherein the video frame consists of an image and  
25 masking; determining (402) an image region of the image within the video frame  
of the video sequence; determining (403) an output aspect ratio; and generating  
(404) an output video frame based on the video frame, the image region, the  
output aspect ratio and a predefined value indicating a filling manner for filling  
the image into the output video frame.

According to an aspect of the present disclosure, it is provided a device for displaying a video sequence, comprising a storage; and a processor for receiving a video frame of the video sequence, wherein the video frame consists of an image and masking; determining an image region of the image within the video frame of the video sequence; determining an output aspect ratio; generating an output video frame based on the video frame, the image region, the output aspect ratio and a predefined value indicating a filling manner for filling the image into the output video frame, wherein the predefined value is stored in the storage.

According to an aspect of the present disclosure, it is provided computer program product downloadable from a communication network and/or recorded on a medium readable by computer and/or executable by a processor, comprising program code instructions for implementing the steps of above method.

According to an aspect of the present disclosure, it is provided non-transitory computer-readable medium comprising a computer program product recorded thereon and capable of being run by a processor, including program code instructions for implementing the steps of above method.

It is to be understood that more aspects and advantages of the disclosure will be found in the following detailed description of the present invention.

### **BRIEF DESCRIPTION OF THE DRAWINGS**

The accompanying drawings, which are included to provide a further understanding of the invention and are incorporated in and constitute a part of this application, will be used to illustrate an embodiment of the invention, as explained by the description. The invention is not limited to the embodiment.

In the drawings:

Fig. 1 is a diagram showing a 2.35:1 widescreen image letterboxed in a 1.33:1 format according to prior art;

Fig. 2 is a diagram showing an image is placed in a wider aspect ratio format according to prior art;

Fig. 3 is a diagram showing an image displayed on TV when the aspect ratio of a video frame having masking is different from aspect ratio of a display device according to prior art;

Fig. 4 is a flowchart showing a method for displaying a video sequence according to an embodiment of the present invention;

Fig. 5 is a diagram showing an image displayed on TV according to the embodiment of the present invention; and

Fig. 6 is a block diagram showing a STB according to the embodiment of the present invention.

### **DETAILED DESCRIPTION**

Sometimes, the aspect ratio of image format for an original image is different from an aspect ratio of a targeting format (or called receiving format), e.g. the aspect ratio of the original image is 4:3 and aspect ratio of the targeting format is 16:9. (the targeting format can be used for storage e.g. in a videotape or hard disk or for displaying on e.g. a TV). In this case, format conversion is necessary. In conventional methods, converting formats of unequal ratios is achieved a) by enlarging the original image to fill the targeting format's display area and cutting off any excess picture information (zooming and cropping), b) by adding horizontal mattes (letterboxing) or vertical mattes (pillarboxing) to retain the original format's aspect ratio for the original image, c) by stretching (hence distorting) the original image to fill the targeting format's aspect ratio, and d) by scaling by different factors in both directions, possibly scaling by a different factor in the center and at the edges (as in Wide Zoom mode).

As to above conventional method b), below gives some description about letterboxing and pillarboxing.

Letterboxing is the practice of transferring a video sequence, e.g. a film shot, with a widescreen aspect ratio to standard-width video formats while preserving

the video sequence's original aspect ratio. The resulting image in the resulting video sequence has mattes (also called black bars or masking) above and below the original image in the standard-width video formats, i.e. the mattes are on the top and bottom of the resulting image. These mattes are part of the resulting image. Fig. 1 is a diagram showing a 2.35:1 widescreen image letterboxed in a 1.33:1 format according to prior art. The rectangle in grid pattern is the original widescreen image and the rectangle including both the rectangle in grid pattern and mattes is the targeting format.

Pillarbox occurs in widescreen video displays when mattes are placed on the sides of the original image. It becomes necessary when video or film that was not originally designed for widescreen is shown on a widescreen displaying device or stored in a widescreen format, or when the original image is displayed or stored in a wider aspect ratio. The original image is placed in the middle of the targeting format. Fig. 2 is a diagram showing an image is placed in a wider aspect ratio format according to prior art.

Sometimes, video sequences received by a set top box (STB) have been processed with letterbox, pillarbox and the like. It means that there is masking around the original image, e.g. as shown in the Fig. 1 and Fig. 2. In this case, the STB can't display the original image (which doesn't have masking) in an appropriate aspect ratio on TV. For example, a Standard Definition (SD) camera shoots a video with aspect ratio of 4:3. The broadcast operator delivers the video in a High Definition channel with aspect ratio of 16:9. In this case, pillarbox is used and the format of video sequence is like what is shown in the Fig. 2. At the user end, if the video sequence is displayed in 16:9, the displayed video is like what is shown in the Fig. 2. But if the video sequence is displayed in 4:3 on TV, it is like what is shown in the Fig. 3. It affects user experience when there is masking around the original image. Similarly, if a HD video is delivered in a SD channel by broadcast operator and the user device displays the received video with aspect ratio of 16:9, the video displayed on TV is like what is shown in the

Fig. 3. In order to solve it, the present disclosure provides a method and a device for display a video sequence.

The embodiment of the present invention will now be described in detail in conjunction with the drawings. In the following description, some detailed  
5 descriptions of known functions and configurations may be omitted for clarity and conciseness.

Fig. 4 is a flowchart showing a method for displaying a video sequence according to an embodiment of the present invention.

In the step 401, the STB receives the video sequence. It is implemented by  
10 tuning to a channel, starting to receive video signal that conveys the video sequence and decoding the video signal to obtain the video sequence. A video frame in the video sequence is made up of the original image and masking around the original image, e.g. as shown in the Fig. 5.

In the step 402, the STB determines image region of the original image in  
15 the video frame. Below gives an example for obtaining image region of the original image. In the example, the video frame is YUV picture in the buffer of the STB. The STB can obtain the size of the video frame, e.g. its width is  $a$  and height is  $b$ , i.e. it has  $b$  rows of pixels and  $a$  columns of pixels. Besides, the STB can obtain data of every pixel in the video frame. Generally, the value for  
20 masking in the video frame is set to a fixed value. The fixed value may vary from implementations, e.g. different chipsets for image processing may use different values. If pixels in a row (or column) have the same value, then it is very likely that the row (or column) belongs to masking. In order to obtain  $n_1$ , the STB scans horizontally from top to bottom. For  $n_2$ , it scans horizontally from bottom  
25 to top. For  $m_1$ , it scans vertically from left to right. And for  $m_2$ , it scans horizontally from right to left. For any scanning for  $n_1$ ,  $n_2$ ,  $m_1$  and  $m_2$ , if pixels in the scanning row have the same value, then it's determined that the scanning row belongs to masking and the scanning continues to next row or column. If they don't have same value, the scanning process stops. In this way, the STB  
30 determines the masking and determines that the real image region starts at  $n_1$

row and  $m_1$  column and ends at  $(b-n_2)$  row and  $(a-m_2)$  column, and the size of the image region is  $(a-m_1-m_2, b-n_1-n_2)$ .

According to a variant, the STB repeats the scanning process for a predetermined number of video frames, e.g. 30 frames to obtain an average values for  $m_1, m_2, n_1, n_2$ .

In the step 403, the STB determines an output aspect ratio. The output aspect ratio can be an input from the user, a predefined value by the user or automatically detected by the STB based on communication protocol between the STB and a displaying device.

In the step 404, the STB generates an output video frame based on the original image, the output aspect ratio and a predefined filling parameter value. The original image is determined by the video frame and the determined image region. The predefined filling parameter indicates a filling manner in which the original image is fitted into the output video frame. Values of filling parameter comprises a) scaling by a same factor (i.e. enlarging or diminishing in a linear manner) the original image to fill entire output video frame and cutting off excess portion of the enlarged image; b) scaling by a same factor the original image so as to make one pair of edges of the original image (e.g. left and right edges, or top and bottom edges) overlap with one pair of edges of the output video frame while keeping the other pair of edges of the original image within the output video frame; c) stretching the original image to fill the entire output video frame; and d) scaling by different factors in the center and at the edges of the original image to fill entire output video frame. The value of the predefined filling parameter is selected by the user from above 4 values and stored in the storage of the STB. The user can use remote to set the predefined filling parameter value in a STB menu. In other words, in this step the STB converts the original image from its original aspect ratio to the output aspect ratio by using a filling manner indicated by the predefined filling parameter value.

It shall note that steps 401 to 404 are performed before the STB starts to display the output video frame to the users. In a variant of the embodiment, it is

performed in a real time manner, i.e. before the STB determines the image region of the original image, the STB displays the video frame of the video sequence to the user, and after the STB determines the image region, the STB displays the output video frame with original image to the user.

5        According to a variant of the embodiment, the STB selects one value close to aspect ratio of the image region from 4:3, 16:9 and 14:9, and uses the selected one as a verified aspect ratio for the image region. So it provides compatibility to existing filling functions among different aspect ratios.

10       Fig. 6 is a block diagram showing a STB according to the embodiment of the present invention.

The STB comprises a tuner, a decoder, a processor and a storage.

The tuner receives a digital broadcasting signal to generate a digital data stream conveying a video sequence, and transmits the digital data stream to the decoder.

15       The decoder, e.g. MPEG-2 decoder, decodes the digital data stream to obtain the video sequence. The obtained video sequence can be either digital signal or analog signal depending on the acceptability of signal to the displaying device, e.g. TV or monitor. If TV only accepts analog signal, then a digital to analog converter shall be added and used.

20       The storage is used to store the predefined filling parameter value.

The buffer (not shown in the Fig. 5) is used to store temporary data, such as video frames of the video sequence, output video frames etc.

25       The processor is used for receiving a video frame of the video sequence, wherein the video frame consists of an image and masking; determining an image region of the image within the video frame of the video sequence; determining an output aspect ratio; generating an output video frame based on the video frame, the image region, the output aspect ratio and a predefined filling parameter value indicating a filling manner for filling the image into the output video frame, wherein the predefined filling parameter value is stored in  
30       the storage.



According to a variant, the processor is further used for receiving a predefined number of video frames; determining image regions for each video frames; and determining average value of the image regions of the predefined number of video frames as a common image region.

5 According to a variant, the processor is further used for determining a value closed to aspect ratio of the image region from 4:3, 16:9 and 14:9 as a verified aspect ratio of the image region.

According to a variant, the processor is further used for performing in the video frame scanning horizontally from top to bottom, scanning horizontally from  
10 bottom to top, scanning vertically from left to right and scanning vertically from right to left; during each scanning, if pixels in a row or column have a same value determining the row or the column belongs to masking; and determining the image region.

According to a variant, the processor is further used for receiving an input  
15 selectable from a) scaling by a same factor the image to fill entire output video frame and cutting off excess portion of the enlarged image; b) scaling by a same factor the image so as to make one pair of edges of the image overlap with one pair of edges of the output video frame while keeping the other pair of edges of the image within the output video frame; c) stretching the image to fill entire output  
20 video frame; and d) scaling by different factors in the center and at the edges of the image to fill entire output video frame; and storing the input as the predefined filling parameter value in the storage.

According to a variant of the embodiment, it is not limited to STB, can be used for a player to play a video sequence stored on a storage device, e.g. local  
25 hard disk or remote hard disk or to play a VOD (Video on Demand) video as long as the video frame consists of an original image (e.g. shot image by a camera) and masking. In this case, the tuner is not necessary.

According to the embodiment of the present invention, it is provided computer program product downloadable from a communication network and/or  
30 recorded on a medium readable by computer and/or executable by a processor,

comprising program code instructions for implementing the steps of above method.

According to the embodiment of the present invention, it is provided non-transitory computer-readable medium comprising a computer program product  
5 recorded thereon and capable of being run by a processor, including program code instructions for implementing the steps of above method.

A number of implementations have been described. Nevertheless, it will be understood that various modifications may be made. For example, elements of different implementations may be combined, supplemented, modified, or  
10 removed to produce other implementations. Additionally, one of ordinary skill will understand that other structures and processes may be substituted for those disclosed and the resulting implementations will perform at least substantially the same function(s), in at least substantially the same way(s), to achieve at least substantially the same result(s) as the implementations disclosed.  
15 Accordingly, these and other implementations are contemplated by this application and are within the scope of the invention as defined by the appended claims.

## CLAIMS

1. A method for displaying a video sequence, comprising the steps of:
  - receiving (401) a video frame of the video sequence, wherein the video
  - 5 frame consists of an image and masking;
  - determining (402) an image region of the image within the video frame of the video sequence;
  - determining (403) an output aspect ratio; and
  - generating (404) an output video frame based on the video frame, the image
  - 10 region, the output aspect ratio and a predefined value indicating a filling manner for filling the image into the output video frame.
  
2. The method of the claim 1, wherein
  - receiving a predefined number of video frames;
  - 15 determining image regions for each video frames; and
  - determining average value of the image regions of the predefined number of video frames as a common image region.
  
3. The method of the claim 1, wherein
  - 20 determining a value closed to aspect ratio of the image region from 4:3, 16:9 and 14:9 as a verified aspect ratio of the image region.
  
4. The method of the claim 1, wherein the step of determining image region further comprises:
  - 25 performing in the video frame scanning horizontally from top to bottom, scanning horizontally from bottom to top, scanning vertically from left to right and scanning vertically from right to left;
  - during each scanning, if pixels in a row or column have a same value
  - determining the row or the column belongs to masking; and
  - 30 determining the image region.

5. The method of the claim 1, wherein

receiving an input for the predefine value selectable from a) scaling by a same factor the image to fill entire output video frame and cutting off excess portion of the enlarged image; b) scaling by a same factor the image so as to make one pair of edges of the image overlap with one pair of edges of the output video frame while keeping the other pair of edges of the image within the output video frame; c) stretching the image to fill entire output video frame; and d) scaling by different factors in the center and at the edges of the image to fill entire output video frame.

10

6. A device for displaying a video sequence, comprising

a storage; and

a processor for receiving a video frame of the video sequence, wherein the video frame consists of an image and masking; determining an image region of the image within the video frame of the video sequence;

15

determining an output aspect ratio; generating an output video frame based on the video frame, the image region, the output aspect ratio and a predefined value indicating a filling manner for filling the image into the output video frame, wherein the predefined value is stored in the storage.

20

7. The device of the claim 6, wherein the processor is further used for

receiving a predefined number of video frames;

determining image regions for each video frames; and

determining average value of the image regions of the predefined number of video frames as a common image region.

25

8. The device of the claim 6, wherein the processor is further used for determining a value closed to aspect ratio of the image region from 4:3, 16:9 and 14:9 as a verified aspect ratio of the image region.

30

9. The device of the claim 6, wherein the processor is further used for performing in the video frame scanning horizontally from top to bottom, scanning horizontally from bottom to top, scanning vertically from left to right and scanning vertically from right to left;

5 during each scanning, if pixels in a row or column have a same value determining the row or the column belongs to masking; and determining the image region.

10. The device of the claim 6, wherein the processor is further used for

10 receiving an input selectable from a) scaling by a same factor the image to fill entire output video frame and cutting off excess portion of the enlarged image; b) scaling by a same factor the image so as to make one pair of edges of the image overlap with one pair of edges of the output video frame while keeping the other pair of edges of the image within the output video frame; c) stretching the image to  
15 fill entire output video frame; and d) scaling by different factors in the center and at the edges of the image to fill entire output video frame; and storing the input as the predefined value in the storage.

11. Computer program product downloadable from a communication network and/or  
20 recorded on a medium readable by computer and/or executable by a processor, comprising program code instructions for implementing the steps of a method according to at least one of claims 1 to 5.

12. Non-transitory computer-readable medium comprising a computer program  
25 product recorded thereon and capable of being run by a processor, including program code instructions for implementing the steps of a method according to at least one of claims 1 to 5.

30

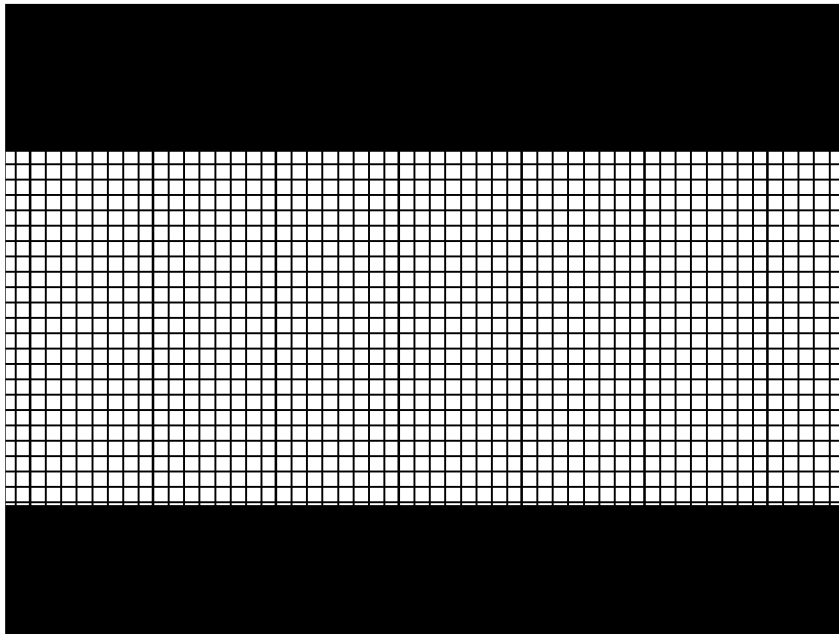


Figure 1 (Prior Art)

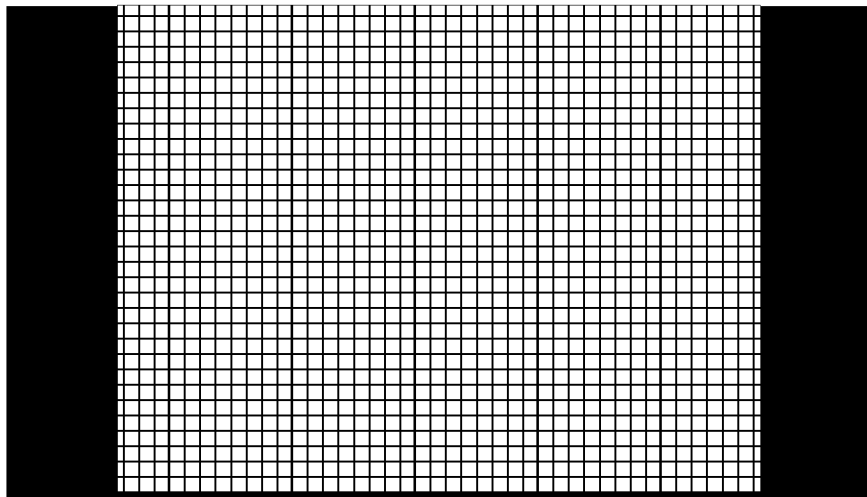


Figure 2 (Prior Art)

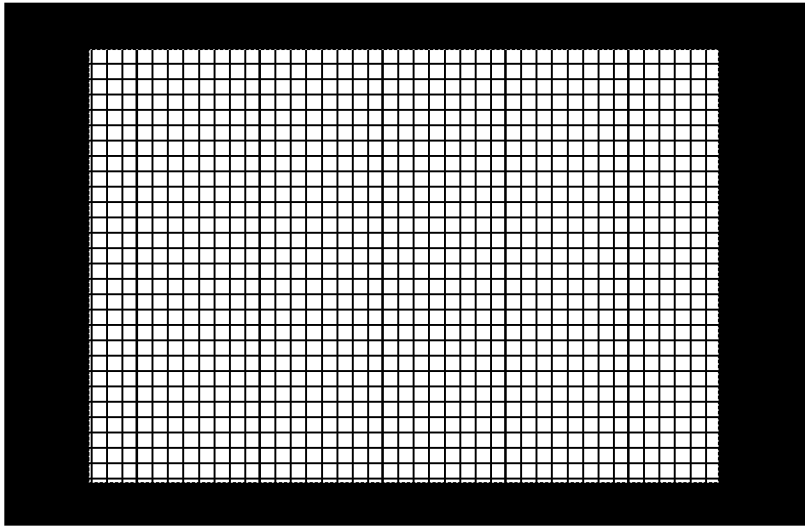


Figure 3 (Prior Art)

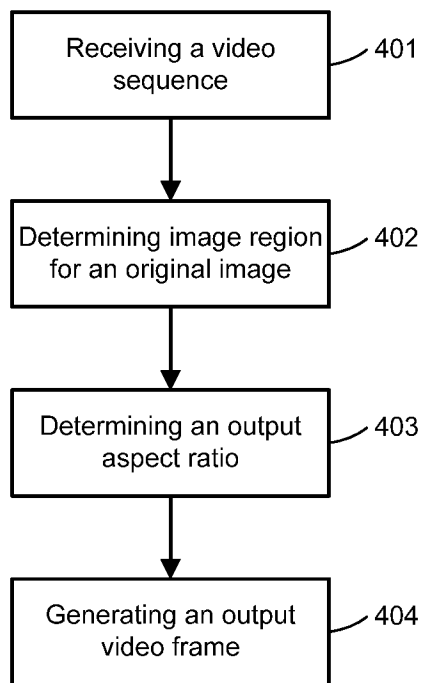


Figure 4

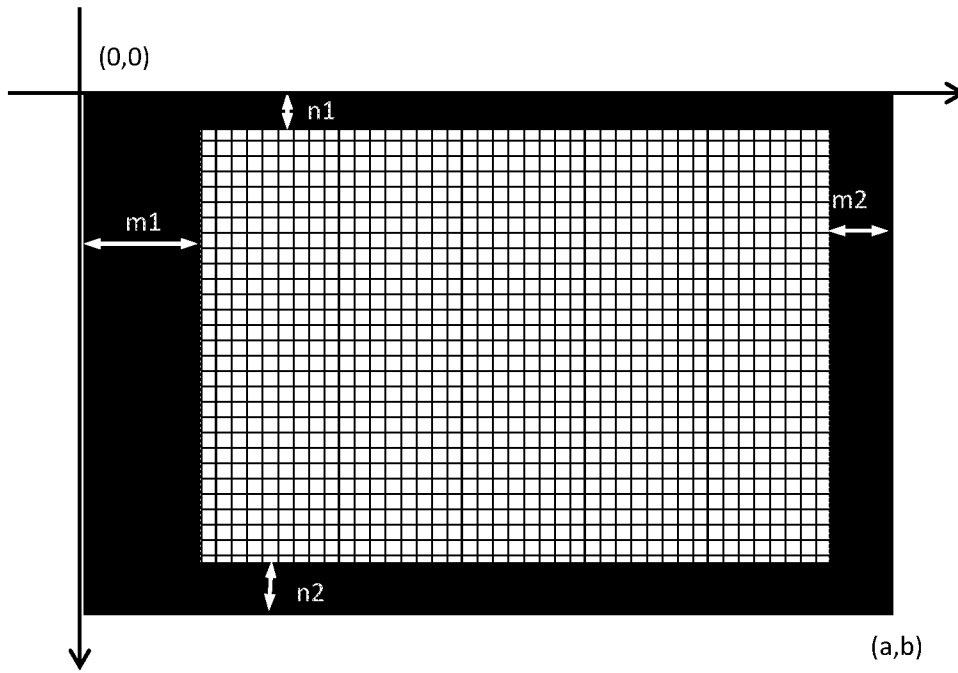


Figure 5

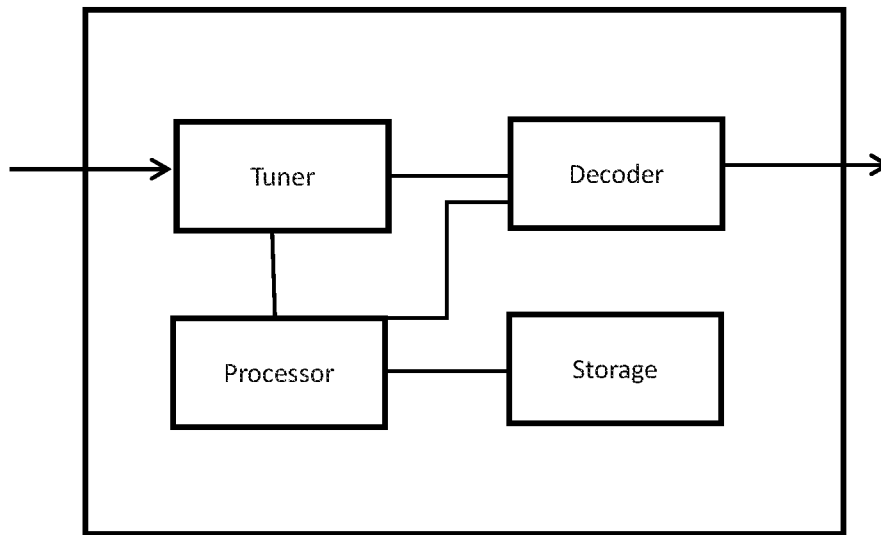


Figure 6



## INTERNATIONAL SEARCH REPORT

International application No.

PCT/CN2014/080979

**A. CLASSIFICATION OF SUBJECT MATTER**

H04N 5/46(2006.01)i

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)

H04N

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)

CNPAT,CNKI,EPODOC,WPI:video,image,frame,masking,letterbox???, pillarbox???,region,portion,area,aspect w ratio,filling

**C. DOCUMENTS CONSIDERED TO BE RELEVANT**

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 2008084503 A1 (SONY CORPORATION) 10 April 2008 (2008-04-10) description, paragraphs [0014], [0086], [0096]-[0100], [0118-0123], [0138], [0146]-[0149] and figures 5, 7A, 8A, 9A	1-3, 5-8, 10-12
A	US 4962521 A (MITSUBISHI DENKI KABUSHIKI KAISHA) 09 October 1990 (1990-10-09) the whole document	1-12
A	US 5208669 A (SONY BROADCAST & COMMUNICATIONS LIMITED) 04 May 1993 (1993-05-04) the whole document	1-12
A	CN 102523390 A (CHENGDU DONGFANG SHENGXING ELECTRONICS CORPORATION) 27 June 2012 (2012-06-27) the whole document	1-12
A	CN 102332096 A (CHINESE ACAD SCI AUTOMATION INST.) 25 January 2012 (2012-01-25) the whole document	1-12

 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

10 March 2015

Date of mailing of the international search report

27 March 2015

Name and mailing address of the ISA/CN

STATE INTELLECTUAL PROPERTY OFFICE OF THE  
P.R.CHINA(ISA/CN)  
6,Xitucheng Rd., Jimen Bridge, Haidian District, Beijing  
100088, China

Authorized officer

FENG,Wei

Facsimile No. (86-10)62019451

Telephone No. (86-10)010-82245514

**INTERNATIONAL SEARCH REPORT**  
**Information on patent family members**

International application No.

**PCT/CN2014/080979**

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)			Publication date (day/month/year)
US	2008084503	A1	10 April 2008	JP	2008098800	A	24 April 2008
US	4962521	A	09 October 1990	EP	0320828	A2	21 June 1989
				DE	3853713	T2	18 January 1996
				KR	920001005	B1	01 February 1992
				JP	H01270475	A	27 October 1989
				JP	H01160285	A	23 June 1989
				JP	H01270474	A	27 October 1989
				JP	H01270473	A	27 October 1989
US	5208669	A	04 May 1993	JP	H05137124	A	01 June 1993
				GB	2251755	A	15 July 1992
CN	102523390	A	27 June 2012			Non	e
CN	102332096	A	25 January 2012			Non	e