

(43) International Publication Date
8 November 2012 (08.11.2012)(51) International Patent Classification:
H04N 13/00 (2006.01)(21) International Application Number:
PCT/EP2012/056063(22) International Filing Date:
3 April 2012 (03.04.2012)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
11305519.8 2 May 2011 (02.05.2011) EP(71) Applicant (for all designated States except US): **THOM-SON LICENSING** [—/FR]; Technicolor, 1-5 rue Jeanne d'Arc, F-92130 Issy-les-Moulineaux (FR).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **VERDIER, Alain** [FR/FR]; Technicolor R&D France, 1 Avenue de Belle Fontaine, F-35510 Cesson-Sévigné (FR). **BOREL, Thierry** [FR/FR]; Technicolor R&D France, 1 Avenue de Belle Fontaine, F-35510 Cesson-Sévigné (FR). **ROBERT, Philippe** [FR/FR]; Technicolor R&D France, 1 Avenue de Belle Fontaine, F-35510 Cesson-Sévigné (FR).(74) Agent: **BROWAEYS, Jean-Philippe**; 1 rue Jeanne d'Arc, F-92443 Issy-les-Moulineaux cedex (FR).

(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, 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, KM, 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, 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, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, 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, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: SMART STEREO GRAPHICS INSERTER FOR CONSUMER DEVICES

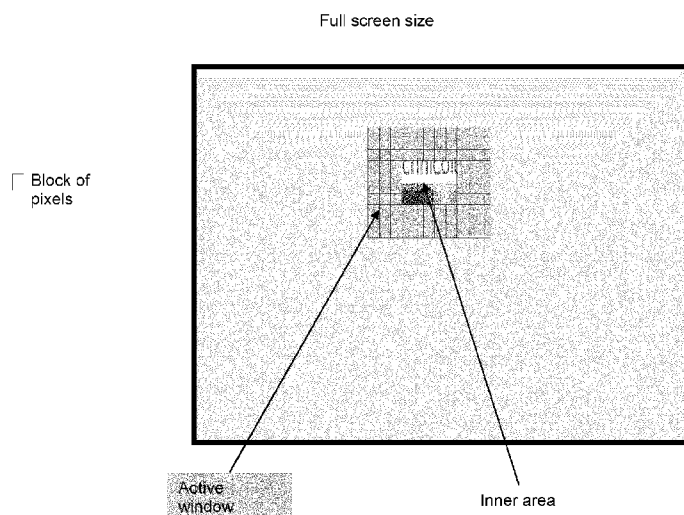


Figure 3

(57) Abstract: The advent of 3D to the home raises a number of challenges beyond the basic display of an available stereo pair. One particular challenge resides in the way to render a graphics element inserted on top of a 3D stereo content at the terminal level. Visual discomfort is likely to occur if graphics is inserted at a fixed depth because of a possible conflict with the surrounding 3D video. This invention introduces a method that aims at removing the depth conflict by ensuring that the graphics element is always perceived at a shorter distance from the viewer than the distance of the occluded video element. The method resides on its capability to being ported on a consumer device, more precisely, the invention describes adaptations at disparity estimation level to fit the limited processing power available in embedded systems.



Declarations under Rule 4.17:

Published:

— *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))* — *with international search report (Art. 21(3))*

SMART STEREO GRAPHICS INSERTER FOR CONSUMER DEVICES

The present invention relates to the display of a three dimensional
(3D) stereo content at the terminal level and more particularly to a method to
5 insert graphics element on top of the 3D stereo content.

3D stereo contents can be watched on a number of various consumer
devices: on 3D TV screens, usually fed by a STB or a Blu-Ray player and
requiring active or passive glasses, but also on mobile devices such as
tablets and smartphone embedding an individual auto-stereoscopic 3D
10 screen. Different kinds of 3D contents can be displayed onto these devices,
either video games if the device is used as a game station, or 3D movies
either shot by stereo cameras (natural sequences) or generated by a
computer (CGI materials). These contents have been built so as to offer to
the viewer the best 3D experience in a standard usage of the display device.
15 As standard usage we mean that only the video content is displayed.

Usually, a graphics overlay is locally inserted on top of the stereo video
content. The most frequent use cases are the insertion of a banner displaying
additional information on the program being watched or a menu to adjust
audio volume or depth level. These latest cases show up intermittent
20 graphics displays but the case of permanent graphics overlay is also possible
in applications where widgets are continuously displayed informing about
stock market, forecast, social network status, etc..

Figure 1 illustrates the case of a video where the main character is
popping out the screen and conflicting with the smiley graphics (bottom left)
25 while the logo inserted on the top right is not conflicting with the surrounding
video (dragon and sky) since located at the screen level. In this example,
both of the graphics are positioned at screen level on the right view as on the
left view. There is no disparity between the right and the left views.

Locating the graphics at a constant minimum disparity will surely also create
30 a visual discomfort after a certain amount of time. Indeed it is painful to have

a graphics constantly close to the viewer, especially if the region of interest is rather located at the screen level or behind.

5 The best way to solve a possible visual conflict is to take into account the disparity of the pixels of the foremost occluded video element so as the graphics is never positioned behind it.

10 As illustrated by Figure 2, the appropriate solution to prevent the visual conflict would have been to shift the smiley rightwards in the left view so as the foremost object in the surrounding of the graphics looks like it is very close to the graphics and just behind (main character's right arm)

A way to facilitate appropriate graphics positioning would be of course, for the stream being decoded, to contain the disparity map together with the video but it is not planned to broadcast more than the minimum and maximum disparity values for the whole program (DVB phase I).

15 Patent application US2011/0018966 discloses a method for adding a caption to a 3D image produced by display patterns displayed on a display screen. The method includes receiving a depth parameter indicative of a frontward location of 3D images, receiving caption data indicative of a caption display pattern and combining the caption data with a subset of the portion of the
20 video content data to create combined pattern data.

However, the presence and position of the graphics overlay is unpredictable and is only dependent upon user preference (widget enabling) or interaction with his device (volume control, overall setting). Thus, at the opposite of the above cited method for adding a caption or a method for sub-titling insertion
25 issue, graphics insertion cannot be solved at the post-production stage.

Thus, according to one aspect of the present invention, a method of inserting a graphic element on a 3D stereo content displayed at a terminal is disclosed. The method is characterised by the steps of:

- determining a block of pixels of the 3D stereo content corresponding
30 to the graphic element;

- extracting dynamically at the terminal level a disparity value (information) for the determined block of the 3D stereo content, indicative of the disparity of the block;
- determining with this extracted disparity value the appropriate depth value for the graphic element to be inserted ;
- and inserting the graphic element at the determined depth value.

With the claimed method, a disparity value for graphics insertion is extracted at the terminal device level which is the level of the consumer device. Thus, the graphics is dynamically positioned at different level according to the surrounding video.

According to another aspect of the invention, the step of extracting dynamically at the terminal level a disparity value (information) for the determined block of the 3D stereo content consists in the steps of

- using a block matching algorithm to select pixels of the determined block,
- determining the corresponding values of the disparity of each selected pixels
- and selecting the minimal value as the disparity value for the determined block of the 3D stereo content (so that the graphic element is always perceived at a shorter distance than the distance of the occluded element).

According to a further aspect of the invention, the block matching algorithm consists in selecting the pixels surrounding the border of the block or in selecting the pixels of subsampling of the block or in selecting the pixels of a search window smaller than the block.

According to a further aspect of the invention, the dynamically extraction of the disparity value is repeated for each image of the 3D stereo content or the dynamically extraction of the disparity value is repeated with a reduced frequency.

According to a further aspect of the invention, the block of pixels of the 3D stereo content corresponding to the graphic element is larger than the size of the graphic element.

5 According to a further aspect of the invention, the appropriate depth value for the graphic element to be inserted corresponds to the addition of an offset value to the said extracted disparity value .

10 According to a further aspect of the invention, a system for inserting a graphic element on a 3D stereo content displayed at a terminal is disclosed.

The system comprises determining means for determining a block of pixels of the 3D stereo content corresponding to the graphic element, extracting means for extracting dynamically at the terminal level a disparity value (information) for the determined block of the 3D stereo content, 15 indicative of the disparity of the block, calculating means for calculating with this extracted disparity value the appropriate depth value for the graphic element to be inserted and inserting means for inserting the graphic element at the determined depth value.

20 According to a further aspect of the invention, the extracting means comprises a block matching algorithm to select pixels of the determined block, determining means to determine the values of the disparity of each selected pixels and selecting means to select the minimal value of the disparity values for the determined block of the 3D stereo content.

25 These and other aspects, features and advantages of the invention will be described or become apparent from the following detailed description of the preferred embodiments, which is to read in connection with the accompanying drawings.

Fig. 1 consists in left and right views showing mispositionned smiley graphics at screen level

Fig. 2 consists in left and right views showing a repositioned graphics

Fig. 3 shows a block for a based local processing

5 Fig. 4 illustrates a process of block matching

Fig. 5 is a graphic showing the offset in function of the disparity

Fig. 6 corresponds to a data flow overview of the invention

10 In the context of finding the appropriate depth for the graphics to be inserted, there are a lot of different methods to extract a disparity map from a stereo pair based on a low complexity algorithm for sparse disparity map extraction. The extraction of the disparity map on a consumer device is clearly a challenging objective: low complexity and good robustness are the key factors to turn a prototype into a product.

15 One of the most common one is based on block matching that consists in finding, for a block of pixel from one view (let's say left view), the displacement to be operated in the alternate view (right view) to match the same block of pixels. This algorithm has good characteristics especially when considering acceleration on parallel architecture.

20 Also it is not the aim to issue a dense disparity map since at the end, only the most negative or minimal disparity value should be retained for graphics depth positioning, a block matching algorithm issuing a disparity value per block is then, far enough in our application. This process is called a "sparse disparity map extraction".

25 Furthermore, a number of characteristics shall be considered in order to reduce the processing time during algorithm execution.

It is the application that decides upon the position of the graphics. So it is not necessary to extract the disparity for the whole picture, only the disparity of

the pixels surrounding the graphics border is required to prevent for the visual conflict.

The following figures illustrate the portion of the incoming video that will be processed during the depth extraction.

- 5 As can be seen on figure 3, the inner pixels of the graphics are not used for the disparity map extraction just because, either they corresponds to a small object that is fully occluded by the graphics or they are part of a larger object that anyway is also present close to the border of the graphics area. This method allows for reducing the number of pixel to be processed, especially
10 for large graphics.

Figure 3 shows a use case where a logo is inserted at the bottom right of the screen. The size of the logo corresponds to 9 X 8 pixels. For example, an active window of 11 pixels to 10 pixels is selected. The inner area is not processed. For example, a border of 3 pixels width highlights the video area
15 that is processed to detect the foremost pixels.

Sub-sampling is also a way to reduce the number of pixel to be processed, it has the associated drawback that it also reduces the precision of the disparity estimation on spatial sub-sampling and temporal resolution on temporal sub-sampling.

- 20 Indeed, if we downsize the incoming video by a factor of N, then performing block matching at one pixel resolution (no sub-pixel) will lead to a disparity value with N pixels precision on original video resolution. As an example, ¼ down-sampling will lead to disparity at +/-4 pixels precision.

Similarly, temporal sub-sampling will induce latency in the disparity
25 estimation update and also possibly a judder effect when moving to various subsequent depth positions.

Furthermore all disparity candidates belonging to the search window will be tested. First, the smaller the search window is, the faster the disparity extraction algorithm is. It is then important to set the search window size
30 appropriately. It is common to consider the extreme values for the disparity at

+/-15% of the picture width. If we consider a full High Definition picture size we obtain +/-96 pixels for minimum and maximum disparity. Obviously, it is preferable to substitute these values to the real minimum and maximum disparity values if they are transported in the stream in compliance with DVB phase I specification. If the real minimum and maximum disparity values are not known, then it is safer to consider +/- 128 pixels as extreme disparity values.

Another characteristic that is important to size appropriately is the precision of the disparity candidate, if 1/N sub-pixel precision is targeted, then the cost will be N times the one at pixel accuracy.

The computing cost can be summarized with the following formula:

$$P = P_0 \cdot \frac{r \cdot f / p}{r_0 \cdot f_0 / p_0}$$

Where:

- P = processeur performance in pixels/picture
- r = « search vector » range in pixels
- p = « search vector » precision in pixel
- f = sub-sampling factor, including both spatial and temporal sub-sampling

with P₀ the processor performance for r₀, p₀ and f₀ arbitrary set to the following values:

- r₀ = 128 pixels
- p₀ = 1 pixel
- f₀ = 1

The formula illustrates the fact that a way to increase the number of pixel processed per second is to reduce the search window, reduce the precision and increase the sub-sampling factor.

Various criteria can be used to check whether a block looks like a similar one. The criterion provides a result that is related to the error of the block matching process. Selection of the best disparity candidate consists in finding the one that minimize the error.

- 5 Figure 4 illustrates the block matching process where a disparity candidate (d) is tested against a matching criterion.

Various criteria can be used to test against similarity, from the most basic expression requiring few processing power to more sophisticated ones but at a higher processing power cost.

- 10 A very simple way to check whether a block looks like another one is the calculation of the _Single Absolute Difference (SAD).

$$SAD(x, y) = \sum_{j=1}^{blockSize} \sum_{i=1}^{blockSize} (|Y_{left}(x + i, y + j) - Y_{right}(x + i + d, y + j)|)$$

Where:

- x, y: coordinates of current block
- 15 • Y left, Y right: luminance value of a pixel in the left and right view
- d: disparity candidate

A slightly higher complexity alternative, in order to cope with unbalanced left and right view to compensate for a possible luminance average difference between left and right views during SAD computation is a weighted SAD:

$$20 \text{ meanLeft}(x, y) = \sum_{j=1}^{blockSize} \sum_{i=1}^{blockSize} (|Y_{left}(x + i, y + j)|)$$

$$\text{meanRight}(x, y) = \sum_{j=1}^{blockSize} \sum_{i=1}^{blockSize} (|Y_{right}(x + i, y + j)|)$$

$$\text{weightedSAD}(x, y) = \sum_{j=1}^{blockSize} \sum_{i=1}^{blockSize} (|Y_{left}(x + i, y + j) - k(x, y, d) * Y_{right}(x + i + d, y + j)|)$$

$$\text{with } k(x, y, d) = \frac{\text{meanLeft}(x, y)}{\text{meanRight}(x + d, y)}$$

- 25 A more complex alternative that takes into account the variance in the block for a better pattern match is the zero-mean Normalized Cross Correlation ZNCC.

$$\text{varLeft}(x, y) = \sum_{j=1}^{blockSize} \sum_{i=1}^{blockSize} (Y_{left}(x + i, y + j) - \text{meanLeft}(x, y))^2$$

$$\text{varRight}(x,y) = \sum_{j=1}^{\text{blockSize}} \sum_{i=1}^{\text{blockSize}} (\text{Yright}(x+i,y+j) - \text{meanRight}(x,y))^2$$

$$\text{covar}(x,y) = \sum_{j=1}^{\text{blockSize}} \sum_{i=1}^{\text{blockSize}} (\text{Yleft}(x+i,y+j) - \text{meanLeft}(x+d,y)), (\text{Yright}(x+i+d,y+j) - \text{meanRight}(x+d,y))$$

$$\text{ZNCC}(x,y) = 1 - \frac{|\text{covar}(x,y)|}{\sqrt{\text{varLeft}(x,y) \cdot \text{varRight}(x,y)}}$$

- 5 Robustness is a very important feature. The invention wants to prevent wrong values to be issued. As only one value is required per frame period, it is clearly worth to add complexity in the domain of confidence assessment and noise reduction rather than on precision and density of the disparity map.

The literature provides a number of techniques to detect areas having erroneous disparity values either because of homogeneous nature of the texture or of an occlusion. Given the limited processing power in a consumer device environment, the invention will focus here on some of the most cost effective tricks to get a first order confidence level:

- 15 - to get a first order confidence level, homogeneous areas should be avoided:

$$G(x,y) = \sum_{j=1}^{\text{blockSize}} \sum_{i=1}^{\text{blockSize}} (|Y(x+i+1,y+j) - Y(x+i,y+j)|)$$

In the case where $G(x,y) < \text{Threshold value}$, then the block matching step for this block is skipped. This has two advantages, the first is to prevent erroneous values to be issued and the second is to speed-up the processing.

- 20 It is well known that homogeneous areas cannot be estimated properly simply because there is no pattern to compare to. A way to skip these blocks is first to compute the gradient of this block and decide whether or not it is worth to enter the block matching algorithm. As the disparity between left and right view is purely horizontal, gradient shall only be computed in the horizontal direction.

- 25 - To get a first order confidence level, a confidence value for the disparity estimate is to be found.

It is difficult to get the right disparity of the background close to the objects border just because it corresponds to an area that is occluded in the alternate

view. It is very likely that the SAD of the best disparity candidate is quite high in this area. a solution is to validate the choice of the best disparity candidate with a comparison to a threshold.

5 A simple way to assess the confidence of the disparity estimate is to compare the SAD to a threshold.

- To get a first order confidence level, appropriate block size is to be determined.

10 As we are not targeting dense disparity map, we have some flexibility to choose the block size. The larger the block is, the more robust the disparity candidate selection is, nevertheless, there is a limit because if we set a large block size, we may never find a correct match and miss small elements in the picture. It has been found that a reasonable block size is 8x8 on a x4 sub-sampled video.

- To get a first order confidence level, isolated disparity values is to be removed

15 It is believed that isolated values are likely to be erroneous values, in order to remove these values a median filter is applied both vertically and temporally. Note that it is not advised to perform horizontal median filtering since vertical edges of objects are sometimes the only part of homogeneous objects where the disparity estimation can be done (presence of horizontal gradient).

- To get a first order confidence level, margin is to be added

25 It is not reasonable to place the graphics exactly at the same depth of the foremost element, just because of the latency: due to the temporal filtering the value provided to the OSD offset manager might be in late in case of decreasing values of the foremost element disparity.

When inserting the graphics, there are different scenarii to reduce the risk of perception conflict:

- First, an offset can be added to systematically shift forward the graphics element.

- Secondly, the strategy can be adapted depending upon the foremost element is in front or behind the screen.
- A solution could be to always put the graphics in front of screen
- An alternate solution could also be to attenuate the disparity when the
5 foremost element is behind the screen or even saturate so as to never render the graphics behind the screen.

Figure 5 correspond to a graphic illustrating these different scenarios. The linear representation of the depth value in function of the minimal disparity value represents the different scenarii.

10 The nr.1 line corresponds to the exact tracking of the foremost object.

The nr.2 line corresponds to the tracking of the foremost object with an additional offset for the security margin.

The nr.3 line shows an alternate solution that differs from the previous one by the fact that the depth reflects only half of the min disparity values behind the
15 screen and the complete before the screen

The nr.4 line shows another alternate solution where the graphics never go behind the screen. The depth will be null for a positive minimal disparity value

The diagram of figure 6 illustrates how the algorithm has been integrated in a
20 consumer device environment (Set Top Box).

The video display pipe comprises:

- a video pipe where the compressed video is decoded, then stored in the video buffer pool and further read for the final display process upon "picture to be displayed" event
- a video buffer pool that contains at least three video buffers. The
25 video data is mapped onto a cacheable external memory area, as the algorithm is accessing several times the same data and the search window is horizontal only, then the cache hit rate is relatively high (>95%). Indeed, if we consider a spatial sub-sampling of 4 on side by

side video material, 8x8 block size and +/-128 pixels search range (on full HD grid) then the corresponding cache size is $1/4 * 1/2 * 8 * 2 * 128 = 256$ bytes which is very low with respect to total available cache size on our target (32 kilobytes).

- 5 • a rescaler that decimates the oldest video buffer to feed the “Sparse Disparity Estimation” block. The rescaler provides the sub-sampled video to the sparse disparity estimator algorithm which calculates the minimum disparity value present in the OSD area so as to manage accordingly the OSD depth positioning prior to mixing with the
- 10 incoming stereo video pipeline.

The system architecture of the Sparse Disparity Estimator is organized as the following:

- 1) a “windowing” unit that extracts from the incoming rescaled video a useful area according to the OSD size/position information;
- 15 2) a “Gradient Computation” unit that receives the useful video from the “windowing” unit and computes gradient at block level;
- 3) A “Block-Matching Core” unit that executes the block matching operation on the useful area on blocks that have gradient level over a given threshold. This unit provides a disparity value per block;
- 20 4) A “Min disparity Select” unit that issue the minimum disparity value from the collection of disparity value provided by the “Block-Matching core” unit;
- 5) A “Median and Temporal Filtering” unit that receives the per-picture disparity values and process a smooth temporal filtering.

25

The section “sparse disparity estimator algorithm” corresponds to the algorithm that has been ported on the CPU embedded in the consumer device, the rest of the diagram shows hardware resources classically used to operate the video pipe. Following the determination of the “On screen display” (OSD) and of the corresponding “window”, the algorithm determines

30

a disparity value for implementing the OSD and with the section "OSD offset management", the appropriate depth for this OSD is defined.

The OSD offset management receives the disparity value to be used as the OSD offset for the current picture.

- 5 The "Mixing" unit receives the video picture and the graphics pixels to be overlaid.

- Performance will be given in terms of percentage of the maximum CPU processing power and also in terms of CPU cycle for the most computing intensive part of the algorithm. Note that the current implementation is purely
- 10 based on the main CPU, no acceleration (GP-GPU) is used. In addition, the main CPU is controlling various dedicated hardware resource during reception, decoding and rendering of the audio/video stream. In this context we assume the algorithm to take less than 10% of the CPU horse power so as not to alter the other real time tasks.
 - 15 Measurements have been done with a mixed of CGI and natural sequences with a graphics size of 1280x240 pixels representing around 15% of the TV screen surface. The sparse disparity estimator takes 2.88% of the CPU load.

Claims

1. A method of inserting a graphic element on a 3D stereoscopic image including a left and a right views displayed at a terminal characterised by the steps of
- determining a block of pixels of the 3D stereoscopic image corresponding to the size of the graphic element and the place for inserting the graphic element;
 - extracting dynamically at the terminal level a disparity value indicative of the frontward location of the determined block of pixels;
 - determining with this extracted disparity value the appropriate depth value for the graphic element to be inserted ;
 - and inserting the graphic element at the determined depth value.
2. The method of claim 1 characterised in that the step of extracting dynamically at the terminal level a disparity value indicative of the frontward location of the determined block of pixels comprises in the steps of
- selecting pixels of the determined block of pixels in using a block matching algorithm to prevent wrong values to be issued ,
 - determining the values of disparity for the selected pixels
 - and selecting the minimal value of the determined values as the disparity value indicative of the frontward location of the determined block of pixels, so that the graphic element is

always perceived at a shorter distance than the distance of the element occluded by the graphic element.

3. The method of claim 2 characterised in that the step of using a block matching algorithm consists in the calculation of the Single Absolute Difference (SAD) to avoid the similarity of blocks of the right and left views.
4. The method of claim 2 characterised in that the selected pixels correspond to the pixels surrounding at least one of the block border.
5. The method of claim 2 characterised in that the selected pixels correspond to a sub-sampling of the block.
6. The method of claim 2 characterised in that the selected pixels correspond to the pixels of a search window smaller than the block.
7. The method of any one of claims 4 to 6 characterised in that the dynamically extraction of the disparity value is repeated for each frame of the 3D stereo content.
8. The method of any one of claims 4 to 6 characterised in that the dynamically extraction of the disparity value is repeated with a reduced frequency.
9. The method of claim 1 characterised in that the block of pixels of the 3D stereo content corresponding to the graphic element is larger than the size of the graphic element.
10. The method of claim 1 characterised in that the appropriate depth value for the graphic element to be inserted corresponds to the said extracted disparity value summed to an offset value.
11. The method of claim 2 characterised in that the method comprises furthermore a step of validating the value of the disparity for each selected pixels with a first order confidence level.
12. A system for inserting a graphic element on a 3D stereo content displayed at a terminal, the system comprising

- determining means for determining a block of pixels of the 3D stereo content corresponding to the graphic element;
- extracting means for extracting dynamically at the terminal level a disparity value (information) for the determined block of the 3D stereo content, indicative of the disparity of the block;
- calculating means for calculating with this extracted disparity value the appropriate depth value for the graphic element to be inserted ;
- and inserting means for inserting the graphic element at the determined depth value.

13. The system of claim 10, characterized in that the extracting means comprises a block matching algorithm to select pixels of the determined block, determining means to determine the values of the disparity of each selected pixels, selecting means to select the minimal value of the disparity values for the determined block of the 3D stereo content.

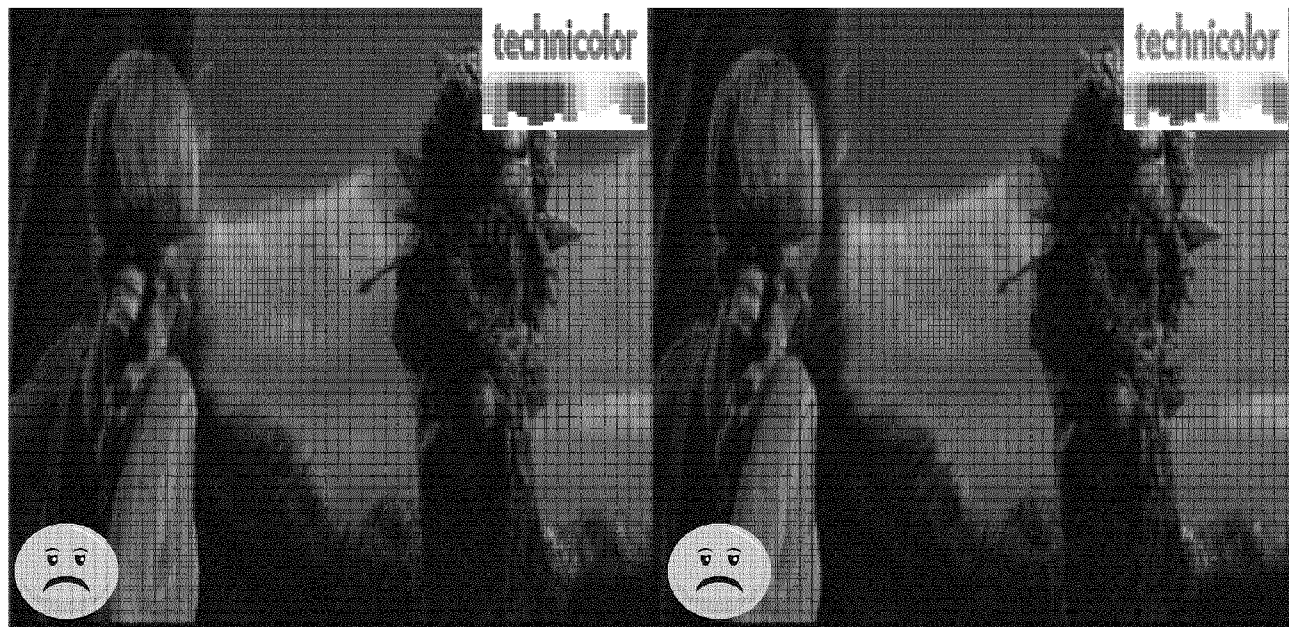


Figure 1

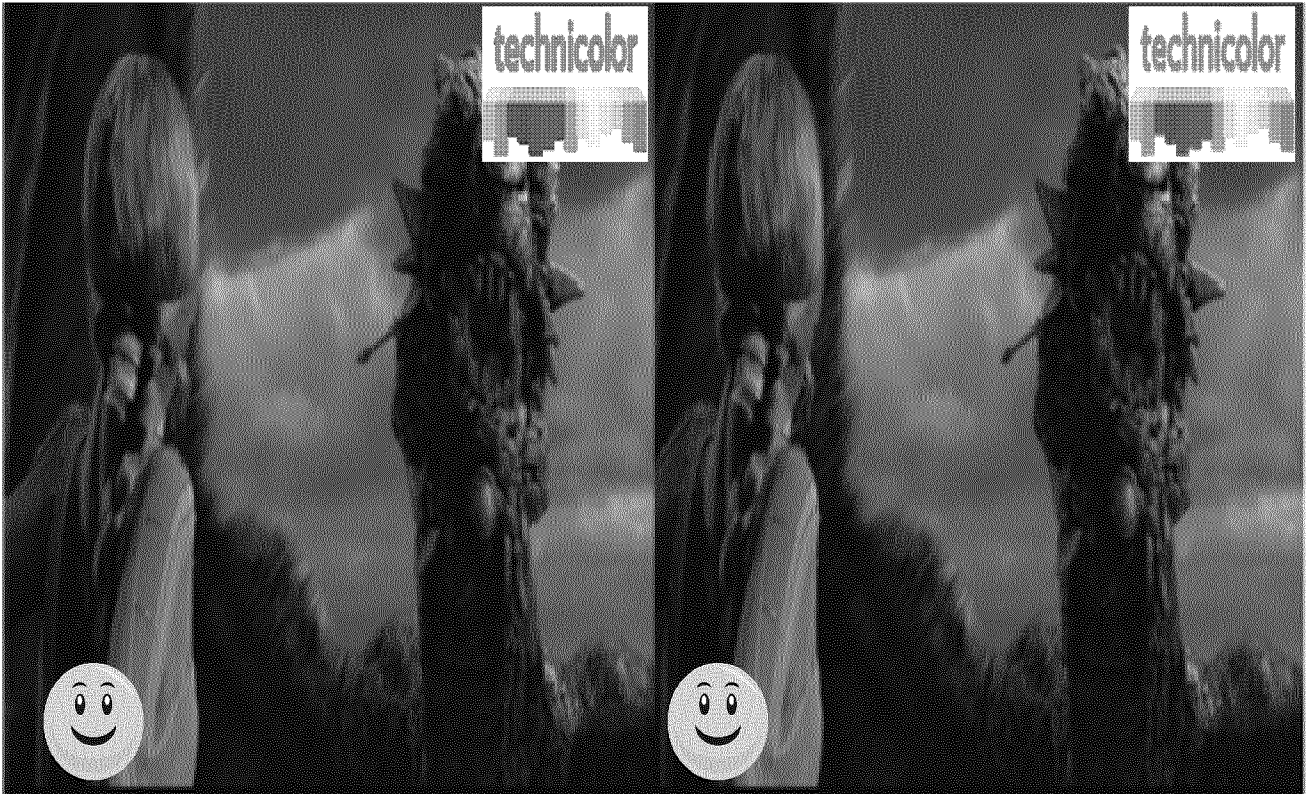


Figure 2

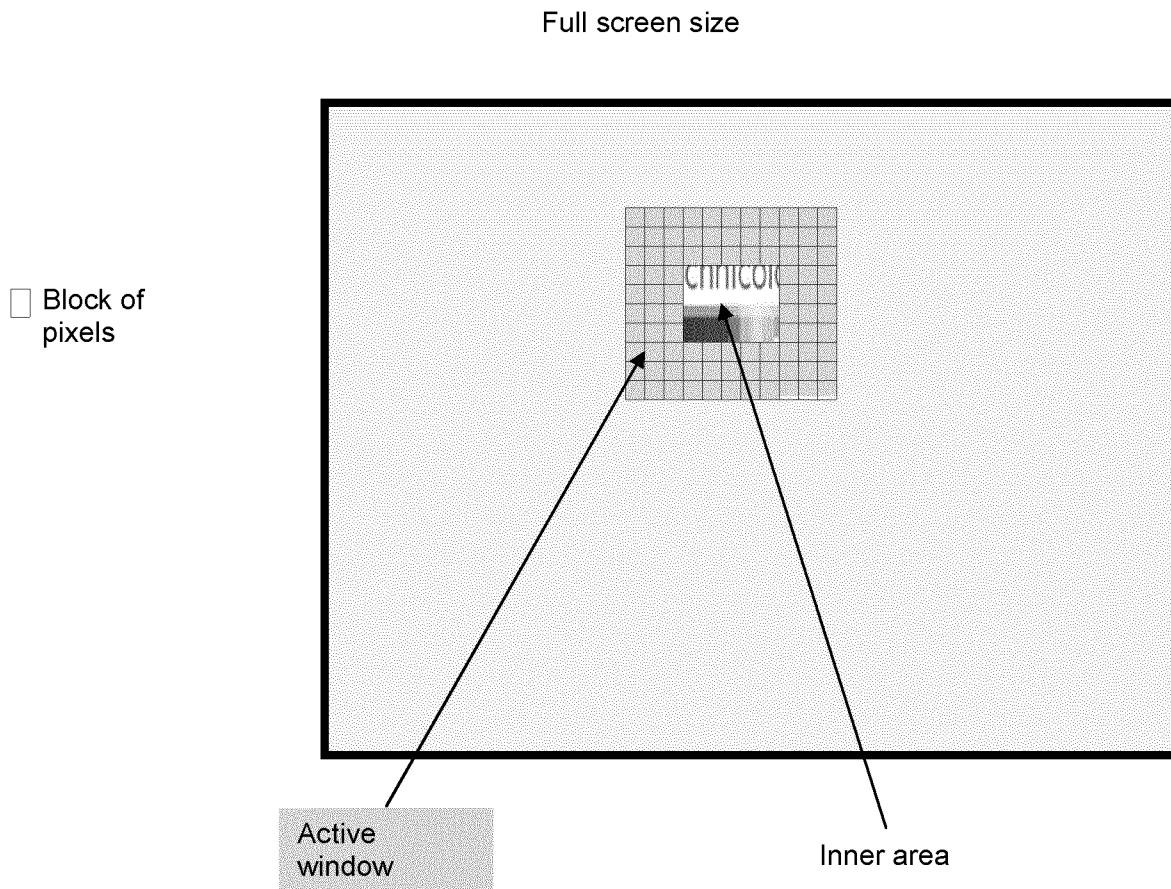


Figure 3

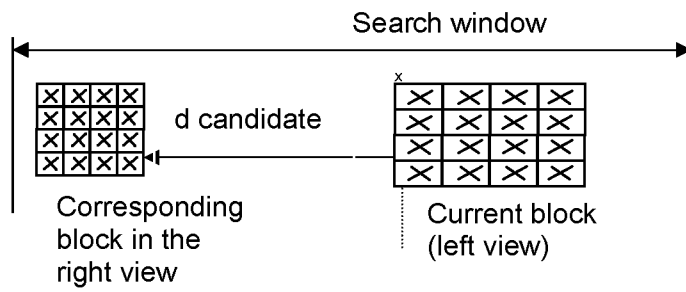


Figure 4

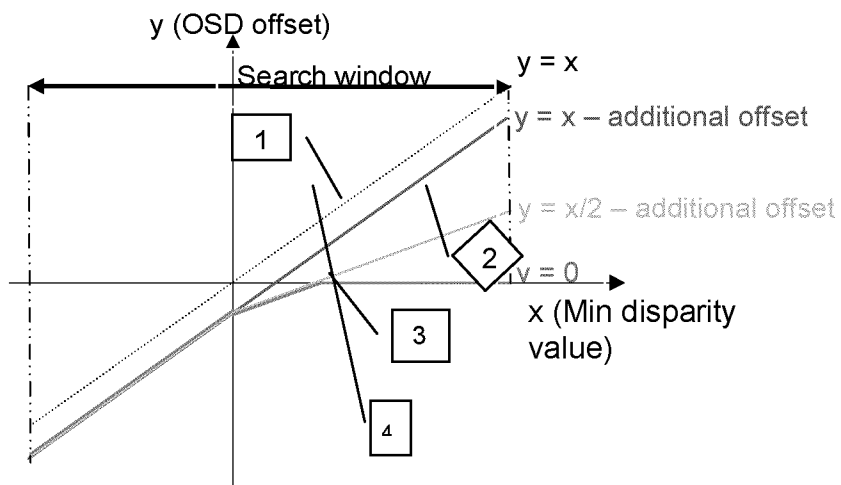


Figure 5

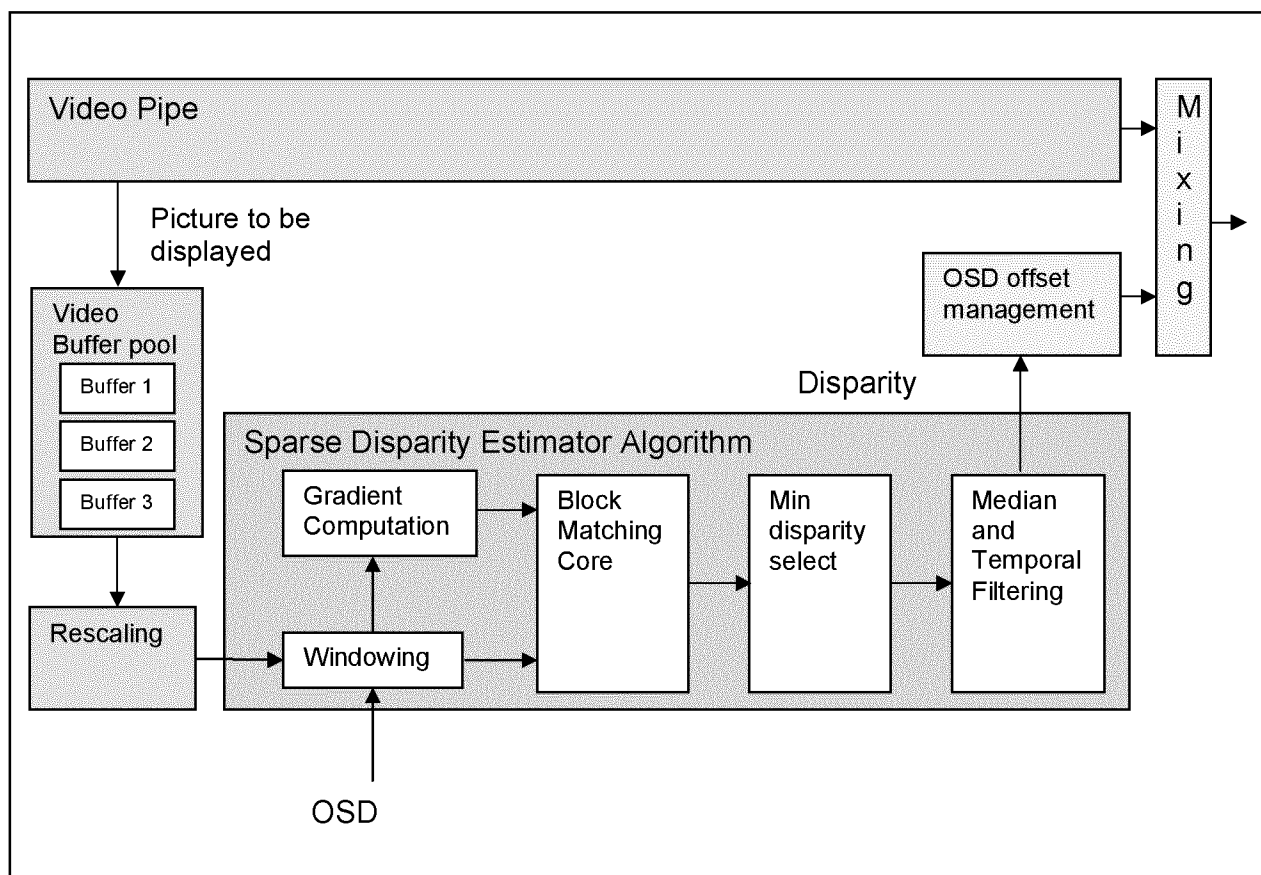


Figure 6

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2012/056063

A. CLASSIFICATION OF SUBJECT MATTER
INV. H04N13/00
ADD.

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)

EP0-Internal, WPI Data

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 2010/095074 A1 (KONINKL PHILIPS ELECTRONICS NV [NL]; PHILIPS INTELLECTUAL PROPERTY [DE] 26 August 2010 (2010-08-26) page 7, line 1 - line 20 page 16, line 1 - line 3 page 12, line 33 - page 13, line 22 page 14, line 21 - line 22 page 10, line 27 - line 32 page 11, line 7 - line 9 -----	1-13
A	GB 2 473 282 A (NDS LTD [GB]) 9 March 2011 (2011-03-09) abstract -----	1-13
A	WO 2008/115222 A1 (THOMSON LICENSING [FR]; IZZAT IZZAT [US]; ZHANG DONG-QING [US]; NIJIM) 25 September 2008 (2008-09-25) abstract -----	1-13
-/-		



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

12 July 2012

Date of mailing of the international search report

23/07/2012

Name and mailing address of the ISA/

European Patent Office, P.B. 5818 Patentlaan 2
NL - 2280 HV Rijswijk
Tel. (+31-70) 340-2040,
Fax: (+31-70) 340-3016

Authorized officer

Cooke, Edward

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2012/056063

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2010/208040 A1 (GUILLOU JEAN-PIERRE [US]) 19 August 2010 (2010-08-19) abstract	1-13
A	----- ATZPADIN N ET AL: "Stereo Analysis by Hybrid Recursive Matching for Real-Time Immersive Video Conferencing", IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE SERVICE CENTER, PISCATAWAY, NJ, US, vol. 14, no. 3, 1 March 2004 (2004-03-01), pages 321-334, XP011108798, ISSN: 1051-8215, DOI: 10.1109/TCSVT.2004.823391 abstract -----	2,13

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2012/056063

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 2010095074 A1	26-08-2010	AU 2010215135 A1 CN 102318352 A EP 2399398 A1 KR 20110126703 A TW 201043002 A TW 201044315 A US 2011304691 A1 WO 2010095074 A1 WO 2010095080 A1	06-10-2011 11-01-2012 28-12-2011 23-11-2011 01-12-2010 16-12-2010 15-12-2011 26-08-2010 26-08-2010
GB 2473282 A	09-03-2011	EP 2462736 A1 GB 2473282 A KR 20120039767 A WO 2011030234 A1	13-06-2012 09-03-2011 25-04-2012 17-03-2011
WO 2008115222 A1	25-09-2008	AT 472230 T CA 2680724 A1 CN 101653011 A EP 2140688 A1 EP 2157803 A1 JP 2010521738 A KR 20090120492 A US 2010238267 A1 WO 2008115222 A1	15-07-2010 25-09-2008 17-02-2010 06-01-2010 24-02-2010 24-06-2010 24-11-2009 23-09-2010 25-09-2008
US 2010208040 A1	19-08-2010	CA 2752497 A1 CN 102326398 A EP 2387852 A2 KR 20110114672 A US 2010208040 A1 WO 2010096281 A2	26-08-2010 18-01-2012 23-11-2011 19-10-2011 19-08-2010 26-08-2010