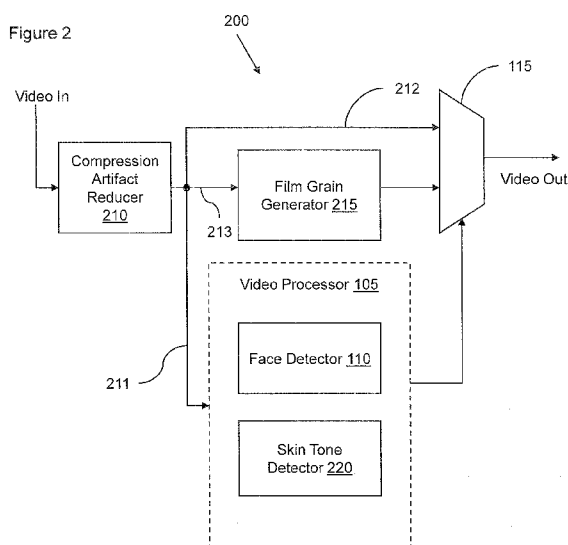




- (51) **International Patent Classification:**
H04N 7/26 (2006.01) *G06T 5/00* (2006.01)
G06K 9/00 (2006.01)
- (21) **International Application Number:**
PCT/US2011/021299
- (22) **International Filing Date:**
14 January 2011 (14.01.2011)
- (25) **Filing Language:** English
- (26) **Publication Language:** English
- (30) **Priority Data:**
61/295,340 15 January 2010 (15.01.2010) US
- (71) **Applicant (for all designated States except US):** **MARVELL WORLD TRADE, LTD** [BB/BB]; L' Horizon, Gunsite Road, Brittons Hill, St Michael 14027 (BB).
- (72) **Inventors; and**
- (75) **Inventors/Applicants (for US only):** **BISWAS, Mainak** [IN/US]; 208 Koshland Way, Santa Cruz, CA 95064 (US). **BALRAM, Nikhil** [IN/US]; 387 Foxborough Drive, Mountain View, CA 94041 (US).
- (74) **Agent:** **KRAGULJAC, Petar**; Kraguljac & Kalnay, LLP, 4700 Rockside Road, Summit One - Suite 510, Independence, OH 44131 (US).
- (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, RO, RS, RU, 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, 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).
- Published:**
— with international search report (Art. 21(3))

(54) **Title:** USE OF FILM GRAIN TO MASK COMPRESSION ARTIFACTS



(57) **Abstract:** Systems, methods, and other embodiments associated with processing video data are described. According to one embodiment, a device comprises a video processor (105) for processing a digital video stream by at least identifying a facial boundary within images of the digital video stream. A combiner (115) selectively applies a digital film grain to the images based on the facial boundary.

USE OF FILM GRAIN TO MASK COMPRESSION ARTIFACTS

CROSS REFERENCE TO RELATED APPLICATIONS

[0001] This application claims the benefit of U.S. provisional application serial
5 No. 61/295,340 filed on January 15, 2010, which is hereby wholly incorporated by
reference.

BACKGROUND

[0002] Bandwidth limitations in storage devices and/or communication
10 channels require that video data be compressed. Compressing video data
contributes to the loss of detail and texture in images. The higher the compression
rate, the more content is removed from the video. For example, the amount of
memory required to store an uncompressed 90-minute long moving picture feature
film (e.g. a movie) is often around 90 Gigabytes. However, DVD media typically
15 has a storage capacity of 4.7 Gigabytes. Accordingly, storing the complete movie
onto a single DVD requires high compression ratios of the order of 20:1. The data
is further compressed to accommodate audio on the same storage media. By
using the MPEG2 compression standard, for example, it is possible to achieve the

relatively high compression ratios. However, when the movie is decoded and played back, compression artifacts like blockiness and mosquito noise are often visible. Numerous types of spatial and temporal artifacts are characteristic of transformed compressed digital video (i.e., MPEG-2, MPEG-4, VC-1, WM9, DIVX, etc.). Artifacts can include contouring (particularly noticeable in smooth luminance or chrominance regions), blockiness, mosquito noise, motion compensation and prediction artifacts, temporal beating, and ringing artifacts.

[0003] After decompression, the output of certain decoded blocks makes surrounding pixels appear averaged together and look like larger blocks. As display devices and televisions get larger, blocking and other artifacts become more noticeable.

SUMMARY

[0004] In one embodiment, a device comprises a video processor for processing a digital video stream by at least identifying a facial boundary within images of the digital video stream. The device also comprises a combiner to selectively apply a digital film grain to the images based on the facial boundary.

[0005] In one embodiment, an apparatus comprises a film grain generator for generating a digital film grain. A face detector is configured to receive a video data stream and determine a face region from images in the video data stream. A combiner applies the digital film grain to the images in the video data stream within the face region.

[0006] In another embodiment, a method includes processing a digital video stream by at least defining a face region within images of the digital video stream;

and modifying the digital video stream by applying a digital film grain based at least in part on the face region.

BRIEF DESCRIPTION OF THE DRAWINGS

5 **[0007]** The accompanying drawings, which are incorporated in and constitute a part of the specification, illustrate various systems, methods, and other embodiments of the disclosure. It will be appreciated that the illustrated element boundaries (e.g., boxes, groups of boxes, or other shapes) in the figures represent one example of the boundaries. In some examples one element may be designed
10 as multiple elements or that multiple elements may be designed as one element. In some embodiments, an element shown as an internal component of another element may be implemented as an external component and vice versa. Furthermore, elements may not be drawn to scale.

[0008] Figure 1 illustrates one embodiment of an apparatus associated with
15 processing digital video data.

[0009] Figure 2 illustrates another embodiment of the apparatus of Figure 1.

[0010] Figure 3 illustrates one embodiment of a method associated with processing digital video data.

20 DETAILED DESCRIPTION

[0011] In the process of video compression, decompression, and removal of compression artifacts, the video stream can often lose a natural-looking

appearance and instead can acquire a patchy appearance. By adding an amount of film grain (e.g. noise), the video stream can be made to look more natural and more pleasing to a human viewer. Addition of film grain may also provide a more textured look to patchy looking areas of the image. When a video stream goes through extensive compression, it can lose much detail in places where there should be texture such as a human face. Typically, the compression process can cause the image in the facial region to look flat and thus unnatural. Applying a film grain to the facial regions may reduce the unnatural look.

[0012] Illustrated in Figure 1 is one embodiment of an apparatus 100 that is associated with using film grain when processing video signals. As an overview, the apparatus 100 includes a video processor 105 that processes a digital video stream (video In). In this example, it is assumed that the video stream was previously compressed and decompressed prior to reaching the video processor. A face detector 110 analyzes the video stream to identify facial regions in the images of the video. For example, a facial region is an area in an image that corresponds to a human face. A facial boundary may also be determined that defines the perimeter of the facial region. In one embodiment, the perimeter is defined by pixels located along the edges of the facial region. A combiner 115 then selectively applies a film grain to the video stream based on the facial boundary. In other words, the film grain is applied to pixels within the facial boundary (e.g., applied to pixels in the facial region). By adding a film grain, facial regions may appear to look more natural rather than appearing unnaturally flat due to compression artifacts. In one embodiment, the film grain is selectively applied by targeting only facial regions and not applying the film grain to other areas as determined by the facial boundaries/regions identified.

[0013] In some embodiments, the apparatus 100 can be implemented in a video format converter that is used in a television, a blue ray player, or other video display device. The apparatus 100 can also be implemented as part of a video decoder for video playback in a computing device for viewing video downloaded from a network. In some embodiments, the apparatus 100 is implemented as an integrated circuit.

[0014] With reference to Figure 2, another embodiment of an apparatus 200 is shown that includes the video processor 105. The input video stream may first be processed by a compression artifact reducer 210 to reduce compression artifacts that appear in the video images. As stated previously, it is assumed the video stream was previously compressed and decompressed. The video stream is output along signal paths 211, 212, and 213, to the video processor 105, the combiner 115, and a film grain generator 215, respectively. As explained above, the facial boundary generated by the video processor 105 controls the combiner 115 to apply the film grain from the film grain generator 215 to the regions in the video stream within the facial boundary. Of course, multiple facial boundaries may be identified for images that include multiple faces.

[0015] With regard to the compression artifact reducer 210, in one embodiment the compression artifact reducer 210 receives the video data stream in an uncompressed form and modifies the video data stream to reduce at least one type of compression artifact. For example, certain in-loop and post-processing algorithms can be used to reduce blockiness, mosquito noise, and/or other types of compression artifacts. Blocking artifacts are distortion that appears in compressed video signals as abnormally large pixel blocks. Also called "macroblocking," it may occur when a video encoder cannot keep up with the allocated bandwidth. It is typically visible with fast motion sequences or quick scene changes. When using

quantization with block-based coding, as in JPEG-compressed images, several types of artifacts can appear such as ringing, contouring, posterizing, staircase noise along curving edges, blockiness in "busy" regions (sometimes called quilting or checkerboarding), and so on. Thus one or more artifact reducing algorithms can be implemented. The particular details of the artifact reducing algorithm that may be implemented with the compression artifact reducer 210 are beyond the scope of the present disclosure and will not be discussed.

[0016] With continued reference to Figure 2, along with the face detector 110, the video processor 105 includes a skin tone detector 220. In general, the face detector 110 is configured to identify areas that are associated with a human face. For example, certain facial features may be located, if possible, such as eyes, ears, and/or mouth to assist in identifying areas of a face. A bounding box is generated that defines a facial boundary of where the face might be. In one embodiment, preselected tolerances may be used to expand the bounding box certain distances from the identified facial features as is expected from typical human head sizes. The bounding box is not necessarily limited to a box shape but may be a polygon, circle, oval, or other curved or angled edges.

[0017] The skin tone detector 220 performs pixel value comparisons that try to identify pixel values that resemble skin tone colors within the bounding box. For example, preselected hue and saturation values that are associated with known skin tone values can be used to locate skin tones in and around the area of the facial bounding box. In one embodiment, multiple iterations of pixel value comparisons may be performed around the perimeter of the bounding box to modify its edges to more accurately find the boundary of the face. Thus the results from the skin tone detector 220 are combined with the results of the face detector

110 to modify/adjust the bounding box of the facial region. The combined results may provide a better classifier of where a face should be in an image.

[0018] In one embodiment, the combiner 115 then applies a digital film grain to the video stream within areas defined by the facial bounding box. For example, the combiner 115 generates masks values using the film grain that are combined with the pixel values within the facial bounding box. In one embodiment, the combiner 115 is configured to apply the digital film grain to red, green, and blue channels in the video data stream. Areas outside the facial bounding box are bypassed (e.g. film grain is not applied). In this manner, the visual appearance of faces in the video may look more natural and have more texture.

[0019] With continued reference to Figure 2, the film grain generator 215 is configured to generate the digital film grain for application to the video stream. In one embodiment, the film grain is generated dynamically (on-the-fly) based on the current pixel values found in the facial regions. Thus the film grain is correlated with the content of the facial region and is colored (e.g., a skin tone film grain). For example, the film grain is generated using red, green, and blue (RGB) parameters from the facial region and are then modified, adjusted, and/or scaled to produce noise values.

[0020] In one embodiment, the film grain generator 215 is configured to control grain size and the amount of film grain to be added. For example, digital film grain is generated that is two or more pixels wide and has particular color values. The color values may be positive or negative. In general, the film grain generator 215 generates values that represent noise with skin tone values, which are applied to the video data stream within the facial regions.

[0021] In another embodiment, the film grain may be generated independently (randomly) from the video data stream (e.g. not dependent upon current pixel values in the video stream). For example, pre-generated skin tone values may be used as noise and applied as the film grain.

5 **[0022]** In one embodiment, the film grain is generated as noise and is used to visually mask (or hide) video artifacts. In the present case, the noise is applied to facial regions of images as controlled by the facial bounding box determined by the face detector 110. Two reasons to add some type of noise to video for display are to mask digital encoding artifacts, and/or to display film grain as an artistic effect.

10 **[0023]** Film grain noise is considered less structured as compared to structured noise that is characteristic of digital video. By adding some amount of film grain noise, the digital video can be made to look more natural and more pleasing to the human viewer. The digital film grain is used to mask unnatural smooth artifacts in the digital video.

15 **[0024]** With reference to Figure 3, one embodiment of a method 300 is shown that is associated with processing video data as described above. At 305, the method 300 processes a digital video stream. At 310, one or more face regions are determined from the video. In one embodiment, a facial boundary is identified and defined for each face within the image(s) to define the corresponding face
20 region. At 315, the digital video stream is modified by applying film grain to the video data based at least in part on the defined face region (or boundaries). For example, using the face region and/or identified facial boundaries as input, the film grain is applied to pixel values that are within the face region. Various ways to generate the film grain, its size, and color can be performed as described
25 previously. In another embodiment, the facial boundary is adjusted by performing

a skin tone analysis as described previously. In this manner, the area that defines the facial region is adjusted with the film grain.

[0025] Accordingly, the systems and methods described herein use noise values that have the visual property of film grain and apply the noise to facial regions in a digital video. The noise masks unnatural smooth artifacts like "blockiness" and "contouring" that may appear in compressed video. Traditional film generally produces a more aesthetically pleasing look than digital video, even when very high-resolution digital sensors are used. This "film look" has sometimes been described as being more "creamy and soft" in comparison to the more harsh, flat look of digital video. This aesthetically pleasing property of film results (at least in part) from the randomly occurring, continuously moving high frequency film grain as compared to the fixed pixel grid of a digital sensor.

[0026] The following includes definitions of selected terms employed herein. The definitions include various examples and/or forms of components that fall within the scope of a term and that may be used for implementation. The examples are not intended to be limiting. Both singular and plural forms of terms may be within the definitions.

[0027] References to "one embodiment", "an embodiment", "one example", "an example", and so on, indicate that the embodiment(s) or example(s) so described may include a particular feature, structure, characteristic, property, element, or limitation, but that not every embodiment or example necessarily includes that particular feature, structure, characteristic, property, element or limitation. Furthermore, repeated use of the phrase "in one embodiment" does not necessarily refer to the same embodiment, though it may.

[0028] “Logic”, as used herein, includes but is not limited to hardware, firmware, instructions stored on a non-transitory medium or in execution on a machine, and/or combinations of each to perform a function(s) or an action(s), and/or to cause a function or action from another logic, method, and/or system. Logic may include a software controlled microprocessor, a discrete logic (e.g., ASIC), an analog circuit, a digital circuit, a programmed logic device, a memory device containing instructions, and so on. Logic may include one or more gates, combinations of gates, or other circuit components. Where multiple logics are described, it may be possible to incorporate the multiple logics into one physical logic. Similarly, where a single logic is described, it may be possible to distribute that single logic between multiple logics. One or more of the components and functions described herein may be implemented using one or more logic elements.

[0029] While for purposes of simplicity of explanation, illustrated methodologies are shown and described as a series of blocks. The methodologies are not limited by the order of the blocks as some blocks can occur in different orders and/or concurrently with other blocks from that shown and described. Moreover, less than all the illustrated blocks may be used to implement an example methodology. Blocks may be combined or separated into multiple components. Furthermore, additional and/or alternative methodologies can employ additional, not illustrated blocks.

[0030] To the extent that the term “includes” or “including” is employed in the detailed description or the claims, it is intended to be inclusive in a manner similar to the term “comprising” as that term is interpreted when employed as a transitional word in a claim.

[0031] While example systems, methods, and so on have been illustrated by describing examples, and while the examples have been described in considerable detail, it is not the intention of the applicants to restrict or in any way limit the scope of the appended claims to such detail. It is, of course, not possible to describe every conceivable combination of components or methodologies for purposes of describing the systems, methods, and so on described herein. Therefore, the disclosure is not limited to the specific details, the representative apparatus, and illustrative examples shown and described. Thus, this application is intended to embrace alterations, modifications, and variations that fall within the scope of the appended claims.

CLAIMS

What is claimed is:

1. A device (100) comprising:

5 a video processor (105) for processing a digital video stream by at least identifying a facial boundary within images of the digital video stream; and

a combiner (115) to selectively apply a digital film grain to the images based on the facial boundary.

10

2. The device of claim 1, wherein the combiner (115) is configured to apply the digital film grain to red, green, and blue channels in the digital video stream.

3. The device of claim 1, further comprising a film grain generator (215) for
15 generating the digital film grain that is correlated to colors of pixels values within the facial boundary.

4. The device of claim 1, wherein the combiner (115) is configured to modify
20 the images by combining the digital film grain with pixel values that are within the facial boundary, and without applying the digital film grain to areas outside the facial boundary.

5. The device of claim 1, further comprising a film grain generator (215) for generating the digital film grain with a size being greater-than-one pixel wide.

6. The device of claim 1, where the video processor comprises:

a skin tone detector (220) for determining skin tone values from pixels in the images to identify portions of a face that are associated with a facial region; and

a face detector (110) configured to determine the facial boundary, which is a boundary of the facial region, where the facial boundary is adjusted based at least in part on the skin tone values.

7. An apparatus (200), comprising:

a film grain generator (215) for generating a digital film grain;

a face detector (110) configured to receive a video data stream (211) and determine a face region from images in the video data stream; and

a combiner (115) to apply the digital film grain to the images in the video data stream within the face region.

8. The apparatus of claim 7, wherein the apparatus is configured to apply the film grain to red, green, and blue channels in the video data stream.

9. The apparatus of claim 7, wherein the film grain generator is configured to generate the digital film grain using red, green, and blue parameters from the video data stream.

10. The apparatus of claim 7, wherein the film grain generator (215) is configured to generate a mask of noise values that are correlated to pixel values of the video data stream, where the mask represents the digital film grain.

5 11. The apparatus of claim 7, where the face detector (110) is configured to generate a bounding box that represents a boundary of the face region within an image; and

where the combiner (115) applies the digital film grain based on the bounding box.

10

12. The apparatus of claim 7, where the face detector (110) comprises:

a skin tone detector (220) for determining skin tone values from pixels in the images to identify portions of a face; and

where the face detector (110) is configured to determine a boundary of the face region, where the boundary is adjusted based at least in part on the skin tone values.

15

13. The apparatus of claim 7, where the combiner (115) is configured to apply the digital film grain to the images within the face region without applying the digital film grain to areas outside the face region.

20

14. The apparatus of claim 7, further comprising a compression artifact reducer (210) configured to:

receive the video data stream in an uncompressed form;

modify the video data stream to reduce at least one type of compression artifact; and

where the apparatus includes signal paths to output the modified video stream to the film grain generator (215), to the face detector (110), and to the combiner (115).

15. A method, comprising:

processing (305) a digital video stream by at least defining (310) a face region within images of the digital video stream; and

modifying (315) the digital video stream by applying a digital film grain based at least in part on the face region.

16. The method of claim 15, wherein the film grain includes color values that are applied to red, green, and blue channels in the video data stream.

17. The method of claim 15, further comprising generating the digital film grain using skin tone values from pixel values from video data stream that are within the face region.

18. The method of claim 15, where the digital film grain is applied to the images within the face region without applying the digital film grain to areas outside the face region.

19. The method of claim 15, further comprising generating the digital film grain from skin tone color values.

20. The method of claim 15, where defining the face region comprises:
determining skin tone values from pixels in the images to identify portions of
a face; and
- 5 adjusting a boundary of the face region based at least in part on the skin
tone values.

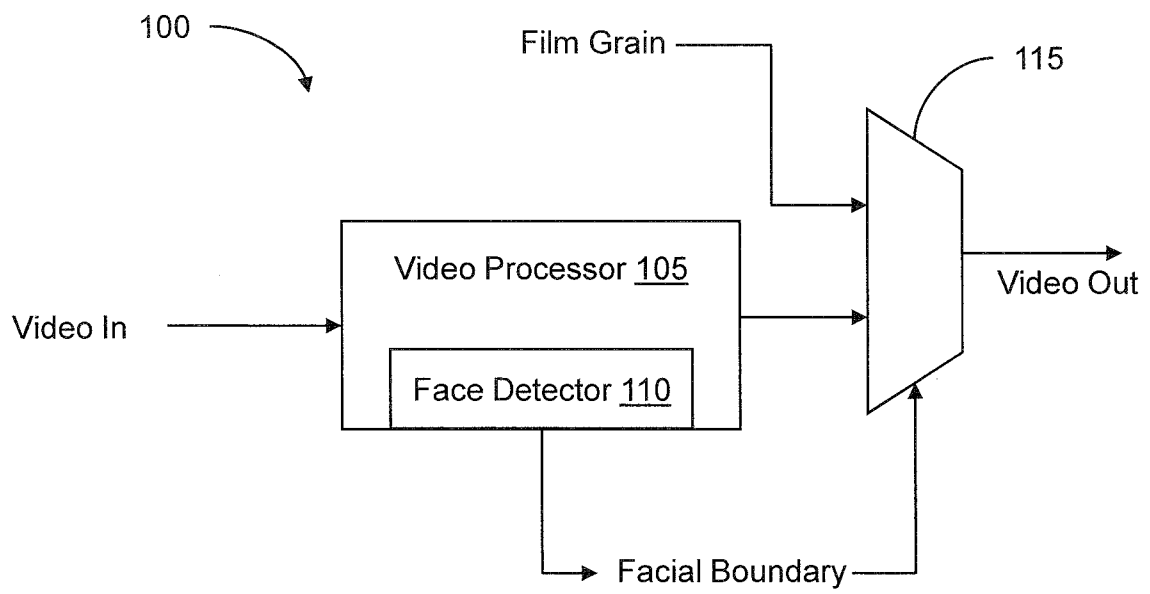


Figure 1

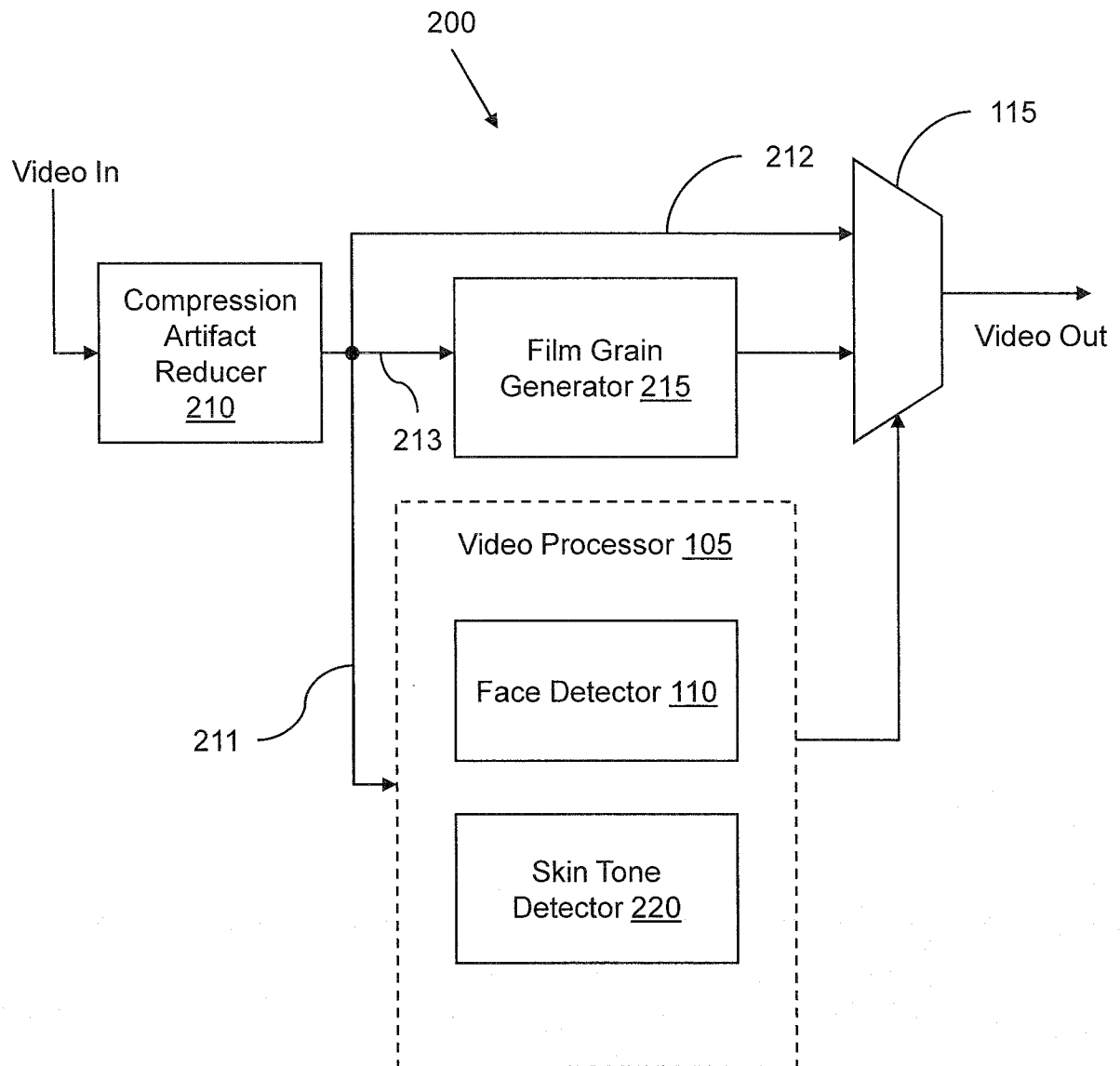


Figure 2

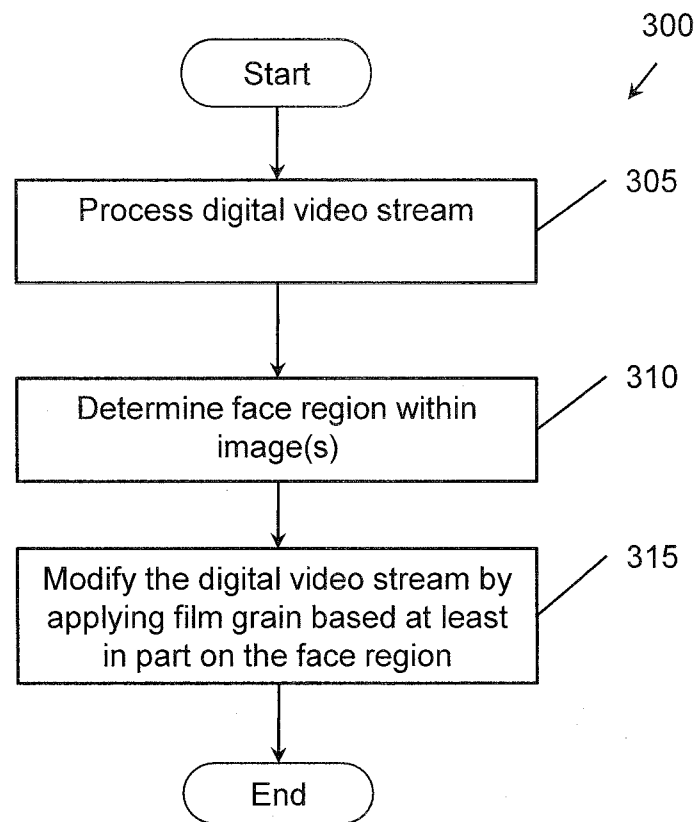


Figure 3

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2011/021299

A. CLASSIFICATION OF SUBJECT MATTER INV. H04N7/26 G06K9/00 G06T5/00 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 G06K 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 practical, search terms used) EPO-Internal, COMPENDEX, INSPEC		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 1 801 751 A2 (MARVELL WORLD TRADE LTD [BB]) 27 June 2007 (2007-06-27) * abstract paragraph [0007] - paragraph [0012]; figure 1 paragraph [0018]; figure 3 -----	1-20
Y	US 2006/182183 A1 (WINGER LOWELL L [CA]) 17 August 2006 (2006-08-17) paragraph [0001] - paragraph [0004] paragraph [0033] paragraph [0045] - paragraph [0055] ----- -/-	1-20
☒ 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 document 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	
28 March 2011	04/04/2011	
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 Colesanti, Carlo	

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2011/021299

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	MENSER B ET AL: "Face detection and tracking for video coding applications", SIGNALS, SYSTEMS AND COMPUTERS, 2000. CONFERENCE RECORD OF THE THIRTY- FOURTH ASILOMAR CONFERENCE ON OCT. 29 - NOV. 1, 2000, PISCATAWAY, NJ, USA, IEEE, vol. 1, 29 October 2000 (2000-10-29), pages 49-53, XP010535333, ISBN: 978-0-7803-6514-8 the whole document	1-20
A	----- DUMITRAS A ET AL: "An automatic method for unequal and omni-directional anisotropic diffusion filtering of video sequences", ACOUSTICS, SPEECH, AND SIGNAL PROCESSING, 2004. PROCEEDINGS. (ICASSP ' 04). IEEE INTERNATIONAL CONFERENCE ON MONTREAL, QUEBEC, CANADA 17-21 MAY 2004, PISCATAWAY, NJ, USA, IEEE, PISCATAWAY, NJ, USA, vol. 3, 17 May 2004 (2004-05-17), pages 317-320, XP010718190, DOI: DOI:10.1109/ICASSP.2004.1326545 ISBN: 978-0-7803-8484-2 the whole document	1-20
A	----- US 2007/236609 A1 (PATHAK BHARAT [IN] ET AL) 11 October 2007 (2007-10-11) paragraph [0058] - paragraph [0059]; table 4 -----	1-20

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/US2011/021299

Patent document cited in search report		Publication date		Patent family member(s)		Publication date
EP 1801751	A2	27-06-2007	CN	101072333 A		14-11-2007
			JP	2007174667 A		05-07-2007
			US	2007140588 A1		21-06-2007
			US	2010166335 A1		01-07-2010

US 2006182183	A1	17-08-2006	US	2009016444 A1		15-01-2009

US 2007236609	A1	11-10-2007	CN	101427560 A		06-05-2009
			EP	2011328 A2		07-01-2009
			JP	2009533898 T		17-09-2009
			KR	20090005163 A		12-01-2009
			WO	2007117623 A2		18-10-2007
