



(51) International Patent Classification:

H04N 7/26 (2006.01) *G06T* 5/00 (2006.01)
H04N 7/50 (2006.01) *G06T* 7/00 (2006.01)

(21) International Application Number:

PCT/EP2010/053211

(22) International Filing Date:

12 March 2010 (12.03.2010)

(25) Filing Language:

English

(26) Publication Language:

English

(30) Priority Data:

09305235.5 13 March 2009 (13.03.2009) EP

(71) Applicant (for all designated States except US): **THOMSON LICENSING** [FR/FR]; 46 Quai Alphonse Le Gallo, F-92100 Boulogne Billancourt (FR).

(72) Inventors; and

(75) Inventors/Applicants (for US only): **GU, Xiaodong** [CN/CN]; 8 Floor, Bldg. A, Technology Fortune Center, No. 8, Xue Qing Rd., Hai Dian District, Beijing, 100085 (CN). **CHEN, Zhibo** [CN/CN]; 8 Floor, Bldg. A, Technology Fortune Center, No. 8, Xue Qing Rd., Hai Dian District, Beijing, 100085 (CN). **LIU, Debeng** [CN/CN]; 8 Floor, Bldg. A, Technology Fortune Center, No. 8, Xue Qing Rd., Hai Dian District, Beijing, 100085 (CN). **XU, Feng** [CN/CN]; 8 Floor, Bldg. A, Technology Fortune Center, No. 8, Xue Qing Rd., Hai Dian District, Beijing, 100085 (CN).

(74) Agent: **RITTNER, Karsten**; Deutsche Thomson OHG, European Patent Operations, Karl-Wiechert-Allee 74, 30625 Hannover (DE).

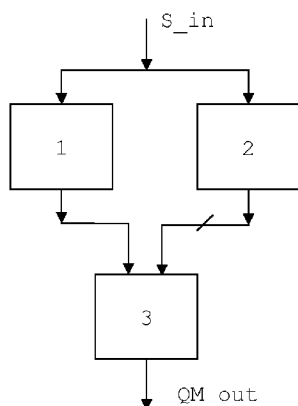
(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, 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, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (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, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

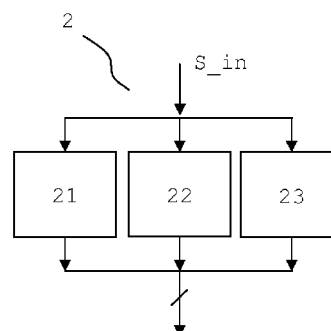
Published:

- with international search report (Art. 21(3))
- before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))

(54) Title: METHOD AND APPARATUS FOR VIDEO QUALITY MEASUREMENT WITHOUT REFERENCE



a)



b)

Fig. 5

(57) Abstract: This invention relates to video quality measurement without having the information of the original signal. The method for measuring video quality of a video received in a data stream comprises steps of extracting from the data stream low level stream information such as QP or bitrate, estimating video quality based on artefacts, comprising determining blur, determining blockiness and determining noise, and combining the low level stream information and the artefacts based estimated video quality, wherein a video quality measure is obtained.

METHOD AND APPARATUS FOR VIDEO QUALITY MEASUREMENT WITHOUT REFERENCEField of the invention

5

This invention relates to a method and apparatus for video quality measurement.

Background

10

This invention relates to video quality measurement without having the information of the original signal. This is called "no-reference" (NR), and it is an essential technique in applications such as video quality monitoring, adaptive coding & streaming, etc.

15

Due to the increasing transmission of digital video contents over broadband and wireless networks, quality monitoring of multimedia data is becoming an important matter. From a quality of experience perspective, it is desirable to evaluate content quality at the receivers. Since the original signals are usually not available at the receiver, quality scores should be provided without having knowledge of the original signals (no-reference metrics, NR) or using very limited information about them. The limited information may be transmitted through a side channel (reduced-reference metrics, RR).

20

25

Engelke and Zepernick¹ provided a detailed analysis of contemporary NR quality metrics. The task of NR quality metrics is very complex, as no information about the original, undistorted signal is available. Therefore, a NR method is an absolute measure of features and properties (such as blockiness and blur) in the distorted signal, which have to

30

¹ U. Engelke, H. Zepernick: "Perceptual-based Quality Metrics for Image and Video Services: A Survey", 3rd EURO-NGI Conference on Next Generation Internet Networks, 2007, p.190-197

be related to perceived quality. And then the perceptual quality prediction is achieved by quantifying different features and combining them in a certain way. The weights for feature quantification are often derived from subjective
5 experiments to find better accordance to perceived quality^{2,3}.

There are recent works^{4,5} on Peak-Signal-to-Noise-Ratio (PSNR) estimation based on the stream information without the presence of the original signal (NR-PSNR). The NR-PSNR
10 is estimated based on the statistical properties of DCT coefficients by supposing the coding parameters from the decoded pictures. Others⁶ take encoding parameters into account, which does not reveal current transmission characteristics.

15

Summary of the Invention

Being a statistical analysis of the video stream, NR-PSNR does not take a Human Vision Model (HVS) into account for
20 video quality estimation. It has been found that it is possible to improve video quality estimation by co-operation between transmitted video streams and reconstructed distorted videos. The artefacts based quality estimation scheme on the reconstructed distorted videos is directly
25 related to the HVS property, and hence more accurate in predicting subjective viewer perception. However, the artefact detection algorithm itself is not so perfect, which leads to an un-steady predicting accuracy. The NR-PSNR

² Z. Wang, H.R. Sheikh, A.C. Bovik: "No-Reference Perceptual Quality Assessment of JPEG Compressed Images", in Proc. of IEEE International Conference on Image Processing, vol.1, Sept.2002, p.477-480

³ M.C.Q. Farias, S.K. Mitra: "No-Reference Video Quality Metric based on Artifact Measurements", in Proc. of IEEE International Conference on Image Processing, vol.3, Sept.2005, p.141-144

⁴ D.S. Turaga, Y. Chen, J. Caviedes: "No-Reference PSNR Estimation for Compressed Pictures", Signal Processing & Image Communication, vol.19, 2004, p.173-184

⁵ A. Ichigaya, M. Kurozumi, N. Hara, Y. Nishida, E. Nakasu: "A Method of Estimating Coding PSNR using Quantized DCT Coefficients", in: IEEE Transactions on Circuits and Systems for Video Technology, vol.16, No.2, Feb.2006

⁶ Zhang et al., US2008/0037864A1

scheme based on the transmitted video streams is much steadier in performance, but lacks correlation with HVS property. In this invention, we provide an improved solution by fusing the two schemes, NR-PSNR and HVS, for better
5 performance. In one aspect of the invention, a method for measuring video quality of a video received in a data stream comprises steps of extracting from the data stream low level stream information such as QP and/or bitrate, estimating video quality based on artefacts, and combining the low
10 level stream information and the artefacts based estimated video quality, wherein a video quality measure is obtained. The estimating comprises determining blur, determining blockiness and determining noise.

15 In another aspect of the invention, an apparatus for measuring video quality of a video received in a data stream comprises extractor means for extracting from the data stream low level stream information such as QP or bitrate, video quality estimator means for estimating video quality
20 based on artefacts, and combiner means for combining the low level stream information and the artefacts based estimated video quality, wherein a video quality measure is obtained. The video quality estimator means, e.g. a processor, and comprises first determining means for determining blur,
25 second determining means for determining blockiness and third determining means for determining noise.

At least the quality estimation accuracy is improved by adding NR-PSNR obtained from low level stream information to the feature space of an artificial neural network for esti-
30 mating quality, together with blur, blockiness and noise. Further objects, features and advantages of the invention will become apparent from a consideration of the following description and the appended claims when taken in connection with the accompanying drawings.

Brief description of the drawings

Exemplary embodiments of the invention are described with reference to the accompanying drawings, which show in

5

Fig.1 distorted images with same mean-square-error (MSE);

Fig.2 a correlation between the prediction accuracy of model FUSE' and the coefficient α ;

Fig.3 an image encoded by H.264 with QP=29;

10

Fig.4 an image encoded by H.264 with QP=37;

Fig.5 a flow chart of the method; and

Fig.6 a block diagram of an apparatus for video quality measurement.

15 Detailed description of the invention**1.1 Artefacts' based Quality Estimation Scheme**

First, an artefacts' based quality estimation scheme is described. Three features are extracted in this scheme:

20

blur, blockiness, and noise.

Image blur is defined as the spread of edges in the spatial domain. If an edge is spread over more pixels (in width), the image is more blurry than if the edge was spread over less pixels. According to the NR blur detection algorithm

25

proposed in the literature⁷, firstly an edge detector (e.g.

vertical Sobel filter) is applied in order to find edges; secondly, each row of the image is scanned and for pixels corresponding to an edge location, the start and end positions of the edge are defined as the local extreme

30

locations closest to the edge, the edge width is given by the difference between the end and start positions, and is identified as the local blur measure for this edge location;

⁷ P. Marziliano, F. Dufaux, S. Winkler, T. Ebrahimi: "A No-Reference Perceptual Blur Metric", in Proc. of IEEE International Conference on Image Processing, 2002, p.57-60

Finally, the global blur measure for the whole image is obtained by averaging the local blur values over all edge locations.

Image blockiness is defined as the ratio between the
5 difference at block boundary and the difference inside the block. Noise measurement is estimated by the intensity variation of the blocks.

A dataset $DS = \{(\text{blur}_i, \text{blockiness}_i, \text{noise}_i, \text{MOS}_i) \mid i=1,2,\dots\}$ is then collected with enough samples, where blur_i , blockiness_i ,
10 and noise_i are the artefacts measure of the sample while MOS_i is the manually labelled quality level.

Since the relationship between human perception and the measurements of the video distortions are difficult to express in a definite manner, a quality prediction model M
15 is built with properly selected machine learning tools, such as Artificial Neural Network (ANN)⁸.

For evaluating the performance of this learning model, a test set is used with extracted features and labelled quality level. E.g. Pearson Correlation (PC) can be adopted
20 for the purpose, defined as the correlation coefficient between model predicted and manually labelled scores, which provides an evaluation of prediction accuracy of the model.

1.2. Statistical based NR-PSNR Scheme

25 Second, a NR-PSNR scheme based on statistics is described. NR-PSNR is proposed based on the below observations: The Mean Square Error (MSE) in the pixel domain is equivalent to the MSE in the DCT domain, because DCT is a normalized orthogonal transformation.
30 When a standard DCT is applied on a natural image or video sequence, the distribution of its AC coefficients shows high

⁸ X.H. Jiang, F. Meng, J.B. Xu, W. Zhou: "No-Reference Perceptual Video Quality Measurement for High Definition Videos based on Artificial Neural Network", in: International Conference on Computer and Electrical Engineering, Dec.2008, p.424-427

regularity. This can be modelled by a Probability Distribution Function (PDF). From such models, Laplacian is the most popular one.

Although the coefficients values have changed after quantization and de-quantization, the distribution trend of the coefficients keeps the same. That is, the distributions of the original DCT coefficients and the de-quantized DCT coefficients can be modelled by a same PDF, as has been proved by some studies.

According to these observations, it is possible to estimate the original DCT coefficients distribution by the quantized DCT coefficients. Therefore, the quantized DCT coefficients are de-quantized, using the QP. The DCT coefficient distribution is measured, since it is the same before quantization and after de-quantization. In the next step, the MSE in the DCT domain caused by quantization can be estimated which is equivalent to the MSE in the pixel domain, and therefore PSNR (NR-PSNR) is estimated.

1.3. Performance Comparison

An artefacts based quality estimation scheme is of much higher correlation with the human perception than statistical based schemes (MSE, PSNR), since it is directly related to the HVS property, as already pointed out in former studies. Fig.1, which shows distorted images with same MSE, is an example from literature⁹: image (a) is the original image, while (b)-(e) are images distorted with same MSE by various kinds of artefacts: b) luminance shift, c) noise, d) blockiness and e) blur respectively.

30

Though images (b)-(e) are of same MSE (thus, PSNR), viewer perception greatly varies.

⁹ Z. Wang, A.C. Bovik: "Mean Squared Error: Love it or Leave it ? – A New Look at Signal Fidelity Measures", IEEE Signal Processing Magazine, vol.26, No.1, Jan.2009, p.98-117

However, such superiority of the artefacts based quality estimation schemes are always abated by the following facts: First, it is impossible to take all kinds of artefacts into account. Those missed components will weaken the superiority since they will be counted in MSE in a steady way. Second, and most important of all, the NR artefacts detection algorithm itself is not so accurate. Taking the algorithms mentioned in section 1.1 as an example, the NR blockiness detection algorithm cannot identify texture on a block boundary, i.e. differ between texture on a block boundary and the blockiness effect; further, the NR blur detection algorithm does not take the source edge width into account.

All these facts resulted in the uncertainty of performance stability of artefacts based quality estimation scheme. On the contrary, statistical based quality estimation schemes (NR-PSNR) have high performance stability while being of much lower correlation with human perception. The present invention provides a way to utilize the strong points of both the two schemes in combination.

2. User study

Results have been verified based on a user study. 30 video shots (all of SD resolution, 720x480 or 720x576, composed by 240 images) are selected in the user study, including movie, cartoon, sports, news, and advertisement, etc. Each video shot is encoded with H.264 to get six processed videos with QP set to 24, 29, 34, 37, 40, and 45 respectively, the GOP structure is IBBP... and GOP size is set to 36.

From all these processed videos, 1288 720x480 images and 1260 720x576 images are randomly selected to generate a dataset $DS = DS_{720 \times 480} \cup DS_{720 \times 576}$ for the user study. The features

(blur, blockiness, noise, NR-PSNR) for each image in DS are extracted according to the method described in Section 1. Ten random subjects (users) are involved in the user study, each is about 3 times the height of TV set away from the 60" display where each image of DS is displayed. The users are asked to give a MOS (Mean Opinion Score) for each displayed image, that is, to identify it into one out of five levels:

Excellent	Good	Fair	Poor	Bad
5	4	3	2	1

Tab.1: MOS levels

10

By average values for all subjects while picking out clear outliers, the MOS score for each image is then defined.

3. Pooling strategies

15 According to the discussion in Sect.1, artefacts based quality estimation is quite intuitive in expressing user experience when browsing videos, and hence should be of high accuracy in estimating viewer perception. However, the estimating accuracy is subdued for some inevitable reasons:

20 first, it is impossible to include all possible artefacts into the scheme, and second the artefact detection algorithm itself is not quite accurate. As a result, the known artefacts based quality estimation scheme is lacking performance stability. On the contrary, NR-PSNR is not expected to

25 provide a very accurate viewer perception since HVS is not considered a statistical based quality estimation scheme. However, NR-PSNR is of high performance stability.

One aspect of the invention is to fuse the two kinds of quality estimation schemes for better performance, better

30 prediction accuracy and/or performance stability. In the following, it is shown that it is advantageous to fuse the two kinds of quality estimation scheme.

3.1. Linear Pooling

Linear pooling is a basic strategy to fuse the artefacts based quality estimation scheme and NR-PSNR:

$$\text{FUSE}' = (1-\alpha) \cdot \text{ANN_Artefacts} + \alpha \cdot \text{NR-PSNR}$$

5 where α is a constant, ANN_Artefacts is the artefacts based quality estimation model, FUSE' is the model to fuse ANN_Artefacts and NR-PSNR with linear pooling strategy. We use dataset DS to train the coefficient α . Fig.2 shows the correlation between the prediction accuracy of model
10 FUSE' (Y axis) and the coefficient α (X axis). Judged by the result, the prediction accuracy reaches its maximum (Pearson Correlation = 0.89, as listed in Tab.2) at about $\alpha=0.3$. That means, the accuracy of prediction is optimized at $\alpha=0.3$. One interesting thing to be noted is that $1-\alpha=0.7$ is nearly the
15 prediction accuracy of the blur and blockiness detection algorithms that were adopted.

3.2. Training by Artificial Neural Network

Generally an ANN-based algorithm includes the following
20 steps: at first, video features are extracted from videos, and subjective assessments are performed on these videos to obtain the MOS score. These features and the corresponding MOS are used for the training step in order to make the ANN learn knowledge of the "distortion" (in our study, 80% of
25 the images in the dataset are used for training and the remaining 20% images are used for testing). After the training, the model will work as a NR model that assesses video quality like human observers in some aspects. In artefacts based quality estimation scheme based on ANN
30 training, the features selected are the appointed artefacts: blur, blockiness, and noise in our study.

In order to fuse the scheme with NR-PSNR, a generalization to this scheme is to make NR-PSNR another extracted feature

in model training. For identification, the model trained with artefacts features is called ANN_Artefacts while the model trained with both, the artefacts features and NR-PSNR, is called FUSE. According to the results reported in Tab.2, the quality estimation accuracy is improved by putting NR-PSNR into the feature space.

	<i>DS</i>	<i>DS</i>_{720x480}	<i>DS</i>_{720x576}
ANN_Artefacts	0.87	0.78	0.86
NR-PSNR	0.74	0.75	0.73
FUSE	0.90	0.86	0.89
FUSE'	0.89	-	-

Tab.2: Pearson Correlation of ANN_Artefacts and FUSE for different dataset

10

Fig.3 and Fig.4 show two examples of quality estimation, featuring the following parameters:

Fig.3: ANN_Artefacts=2.62; NR-PSNR=37.04; FUSE=3.32; MOS=4.0

Fig.4: ANN_Artefacts=2.92; NR-PSNR=33.06; FUSE=2.50; MOS=2.5

15

Fig.3 is selected from a video sequence encoded by H.264 and QP=29 while Fig.4 is selected from a video sequence encoded with QP=37 (QP=Quantization parameter). Fig.3 is of higher user perception, as shown by its higher MOS value. However the artefacts based quality estimation scheme (FUSE) gives similar scores to both the two images, mainly because the blur detection algorithm fails to identify the special shooting technique in creating Fig.3. This led to a more inaccurate blur score. The problem is settled to some extent with the presence of NR-PSNR in FUSE. It is an example to show that the model becomes more stable with the presence of NR-PSNR, especially when the artefact detection algorithm does not give an accurate score.

20

25

4. Conclusion

Two fusion strategies are proposed to improve the performance of artefacts based quality estimation schemes with the help of stream information, in both accuracy and stability. The invention is a pioneer work in adopting stream information to improve the performance of NR quality measurement. Its importance becomes clear considering the fact that stream information is provided in most application scenarios.

10

In the examples, NR-PSNR is selected as the representative of stream information for simplicity. Other low level, basic stream information can be used instead.

Fig.5 shows a flow chart of the method for measuring video quality of a video received in a data stream S_{in} . The first step is extracting 1 from the data stream low level stream information such as QP or bitrate. The second step, which can be done before, after or at the same time as the first step, is estimating 2 video quality QM_2 based on artefacts. This may comprise determining blur 21, determining blockiness 22 and determining noise (PSNR) 23. The third step is combining 3 the low level stream information and the artefacts based estimated video quality, wherein a video quality measure QM_{out} is obtained. Particular embodiments are described below.

In one embodiment of the invention, the low level stream information comprises bitrate or quantization parameter (QP). In one embodiment, the step of combining the low level stream information and the artefacts based estimated video quality comprises steps of calculating a NR-PSNR value from the low level stream information, and calculating a weighted sum of said NR-PSNR value and a value resulting from said

artefacts based estimated video quality. The weighted sum can be calculated according to

$$\text{FUSE}' = (1 - \alpha) \cdot \text{ANN_Artefacts} + \alpha \cdot \text{NR-PSNR}.$$

In one embodiment, the calculating of a NR-PSNR value from
5 low level stream information comprises de-quantizing received DCT coefficients of a block using at least QPs of the low level stream information, calculating a NR-PSNR value from the de-quantized DCT coefficients, determining a distribution of the de-quantized DCT coefficients in the
10 block, and calculating a mean-square-error (MSE) of the determined distribution, wherein the mean-square-error represents said NR-PSNR value.

In one embodiment, a dataset $DS = \{ (\text{NR-PSNR}_i, \text{blur}_i, \text{noise}_i, \text{blockiness}_i, \text{MOS}_i) \mid i=1,2,\dots,N \}$ is collected with N sample
15 images, where blockiness_i , blur_i and noise_i are the artefacts measure of a sample image, NR-PSNR_i is said NR-PSNR value calculated from low level stream information, and MOS_i is a manually labelled quality level of the sample image, and the
20 method further comprises training a machine learning tool with the dataset. In one embodiment, determining blur comprises linewise scanning, detecting edges and measuring the width of the detected edges.

In one embodiment, said measuring the width of the detected
25 edges comprises determining start and end positions of an edge, wherein the width is given by the difference between the start and end positions. In one embodiment, the start and end positions of the detected edge are local extreme locations of luminance that are closest to the edge. In one
30 embodiment, the edge width is used as the local blur measure for a current edge location, further comprising the step of averaging the local blur measures over all edge locations to obtain a global blur measure for the whole image. In one embodiment, the step of determining blockiness comprises

determining a ratio between the difference at block boundary and the difference inside the block. In one embodiment, the step of determining noise comprises determining a luminance intensity variation of a block, e.g. by determining a
5 standard deviation of luminance values within a block.

Fig.6 shows a block diagram of an apparatus 60 for video quality measurement of a video received in a data stream, according to one aspect of the invention. The apparatus 60
10 for measuring video quality comprises extractor means 61 for extracting from the data stream S_in low level stream information, video quality estimator means 62 for estimating video quality QM2 based on artefacts, wherein the video quality estimator means 62 comprises first determining means
15 621 for determining blur, second determining means 622 for determining blockiness and third determining means 623 for determining noise, and first combiner means 63 for combining the low level stream information and the artefacts based estimated video quality, wherein a video quality measure
20 QM_out is obtained. Measures of the determined blur, noise and blockiness can be combined in a second combiner means 624, which results in a combined measure QM2 of blur, noise and blockiness. The second combiner means 624 is optional.

25 In one embodiment, the low level stream information comprises quantization parameter or bitrate. In one embodiment, the first combiner means 63 for combining the low level stream information and the artefacts based estimated video quality comprises first calculator means 631
30 for calculating an NR-PSNR value from the low level stream information, and second calculator means 632 for calculating a weighted sum of said NR-PSNR value and a value (such as ANN_Artefacts) resulting from said artefacts based estimated video quality. In one embodiment, the first calculator means

631 for calculating a NR-PSNR value from said low level stream information comprises a de-quantizer, and further distribution determining means 6311 for determining a distribution of de-quantized DCT coefficients in a block, and MSE calculator means 6312 for calculating a mean-square-error (MSE) of the determined distribution, wherein the mean-square-error is said NR-PSNR value. In one embodiment, the first determining means 621 for determining blur comprises scanning means for linewise scanning an image, an edge detector for detecting edges and a measurement module for measuring the width of the detected edges. In one embodiment, said measurement module determines start and end positions of an edge, and provides the width as being the difference between the start and end positions.

15

In one embodiment, the start and end positions of the detected edge are local extreme locations of luminance that are closest to the edge. In one embodiment, the edge width is used as the local blur measure for a current edge location, and the measurement module further comprises averaging means for averaging the local blur measures over all edge locations, thus obtaining a global blur measure for the whole image. In one embodiment, the second determining means 622 for determining blockiness determines a ratio between the difference at block boundary and the difference inside the block. In one embodiment, the third determining means 623 for determining noise determines a luminance intensity variation within a block.

30 While there has been shown, described, and pointed out fundamental novel features of the present invention as applied to preferred embodiments thereof, it will be understood that various omissions and substitutions and changes in the apparatus and method described, in the form

and details of the corresponding devices and in their operation, may be made by those skilled in the art without departing from the spirit of the present invention. Although the present invention has been disclosed with regard to only
5 certain stream parameters, one skilled in the art would recognize that the method and devices described herein may be applied to any stream parameters. It is expressly intended that all combinations of those elements that perform substantially the same function in substantially the
10 same way to achieve the same results are within the scope of the invention. Substitutions of elements from one described embodiment to another are also fully intended and contemplated.

15 It will be understood that the present invention has been described purely by way of example, and modifications of detail can be made without departing from the scope of the invention.

20 Each feature disclosed in the description and (where appropriate) the claims and drawings may be provided independently or in any appropriate combination. Features may, where appropriate be implemented in hardware, software, or a combination of the two. Connections may, where
25 applicable, be implemented as wireless connections or wired, not necessarily direct or dedicated, connections. Reference numerals appearing in the claims are by way of illustration only and shall have no limiting effect on the scope of the claims.

Claims

1. A method for measuring video quality of a video received in a data stream, comprising steps of
 - 5 - extracting (1) from the data stream low level stream information;
 - estimating (2) video quality based on artefacts, comprising determining blur, determining blockiness and determining noise (PSNR); and
 - 10 - combining (3) the low level stream information and the artefacts based estimated video quality, wherein a video quality measure is obtained.
2. Method according to claim 1, wherein said low level
15 stream information comprises quantization parameter.
3. Method according to claim 1 or 2, wherein the step of combining the low level stream information and the artefacts based estimated video quality comprises
20 steps of
 - calculating a NR-PSNR value from the low level stream information, and
 - calculating a weighted sum of said NR-PSNR value and a value (ANN_Artefacts) resulting from said
25 artefacts based estimated video quality.
4. Method according to one of the claims 1-3, further comprising steps of
 - 30 - de-quantizing received DCT coefficients of a block, using said low level stream information;
 - calculating a NR-PSNR value from the de-quantized DCT coefficients;
 - determining a distribution of the de-quantized DCT coefficients in the block; and

- calculating a mean-square-error (MSE) of the determined distribution, wherein the mean-square-error is said NR-PSNR value.

- 5 5. Method according to claim 3, wherein a dataset $DS = \{ (NR\text{-}PSNR_i, blur_i, blockiness_i, noise_i, MOS_i) \mid i=1,2,\dots,N \}$ is collected with N sample images, where $blockiness_i$, $blur_i$ and $noise_i$ are the artefacts measure of a sample image, $NR\text{-}PSNR_i$ is said NR-PSNR value calculated from
10 low level stream information, and MOS_i is a manually labelled quality level of the sample image, further comprising a step of training a machine learning tool with the dataset.
- 15 6. Method according to one of the claims 1-5, wherein determining blur comprises linewise scanning, detecting edges and measuring the width of the detected edges.
- 20 7. Method according to claim 6, wherein said measuring the width of the detected edges comprises determining start and end positions of an edge, wherein the width is given by the difference between the start and end positions.
- 25 8. Method according to claim 6 or 7, wherein the start and end positions of the detected edge are local extreme locations of luminance that are closest to the edge, and the edge width is used as the local blur
30 measure for a current edge location, further comprising a step of averaging the local blur measures over all edge locations to obtain a global blur measure for the whole image.

9. Method according to one of the claims 1-8, wherein the step of determining blockiness comprises determining a ratio between luminance differences at a block boundary and luminance differences inside the block.

5

10. Method according to one of the claims 1-9, wherein the step of determining noise comprises determining a luminance intensity variation of a block.

10

11. An apparatus for measuring video quality of a video received in a data stream, comprising

- extractor means (61) for extracting from the data stream (S_in) low level stream information;

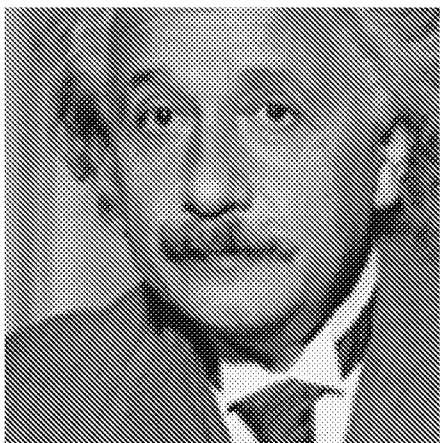
15

- video quality estimator means (62) for estimating video quality based on artefacts, wherein the video quality estimator means comprises first determining means (621) for determining blur, second determining means (622) for determining blockiness and third determining means (623) for determining noise; and

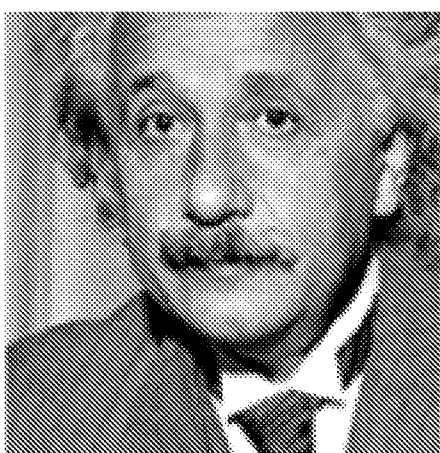
20

- combiner means (63) for combining the low level stream information and the artefacts based estimated video quality, wherein a video quality measure (QM_out) is obtained.

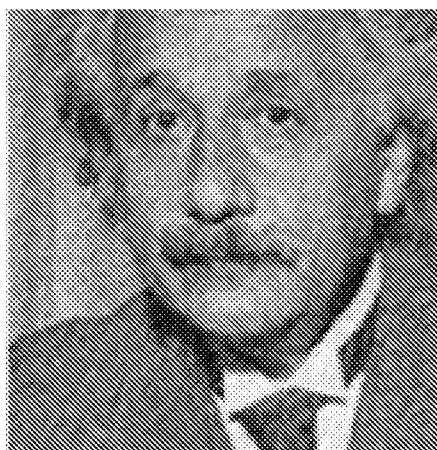
25



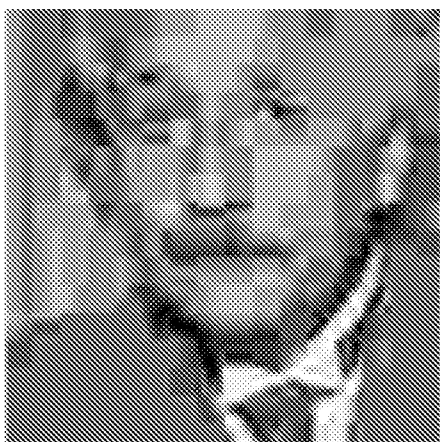
a)



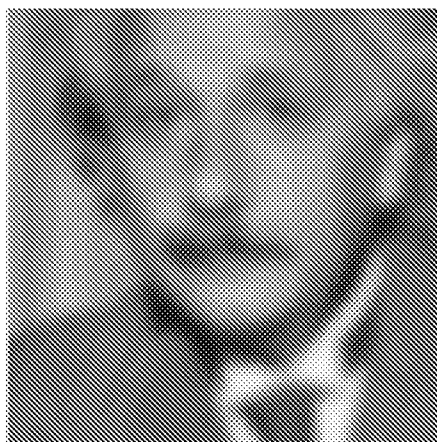
b)



c)



d)



e)

Prior art

Fig.1

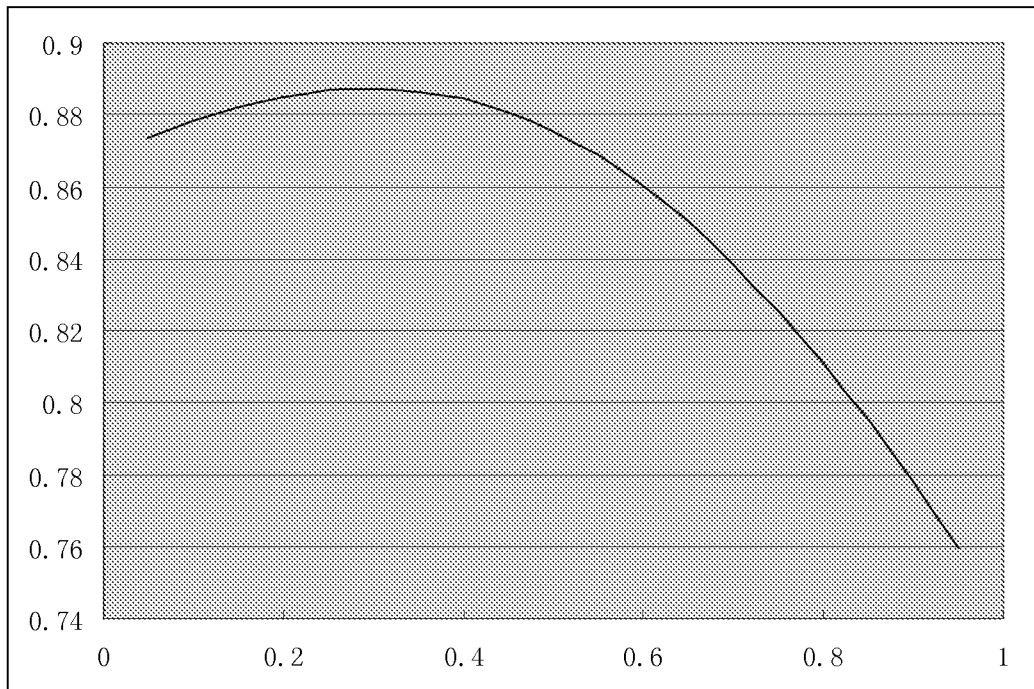
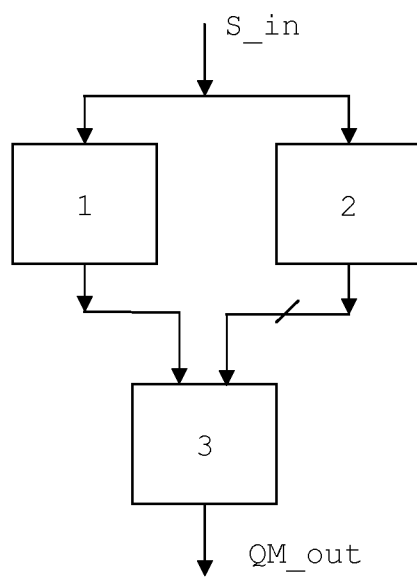
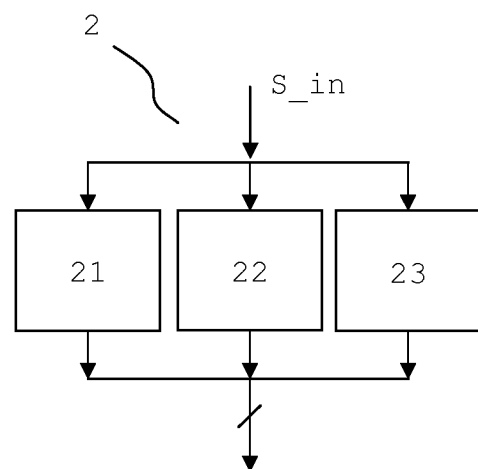
**Fig.2****Fig.3**



Fig.4

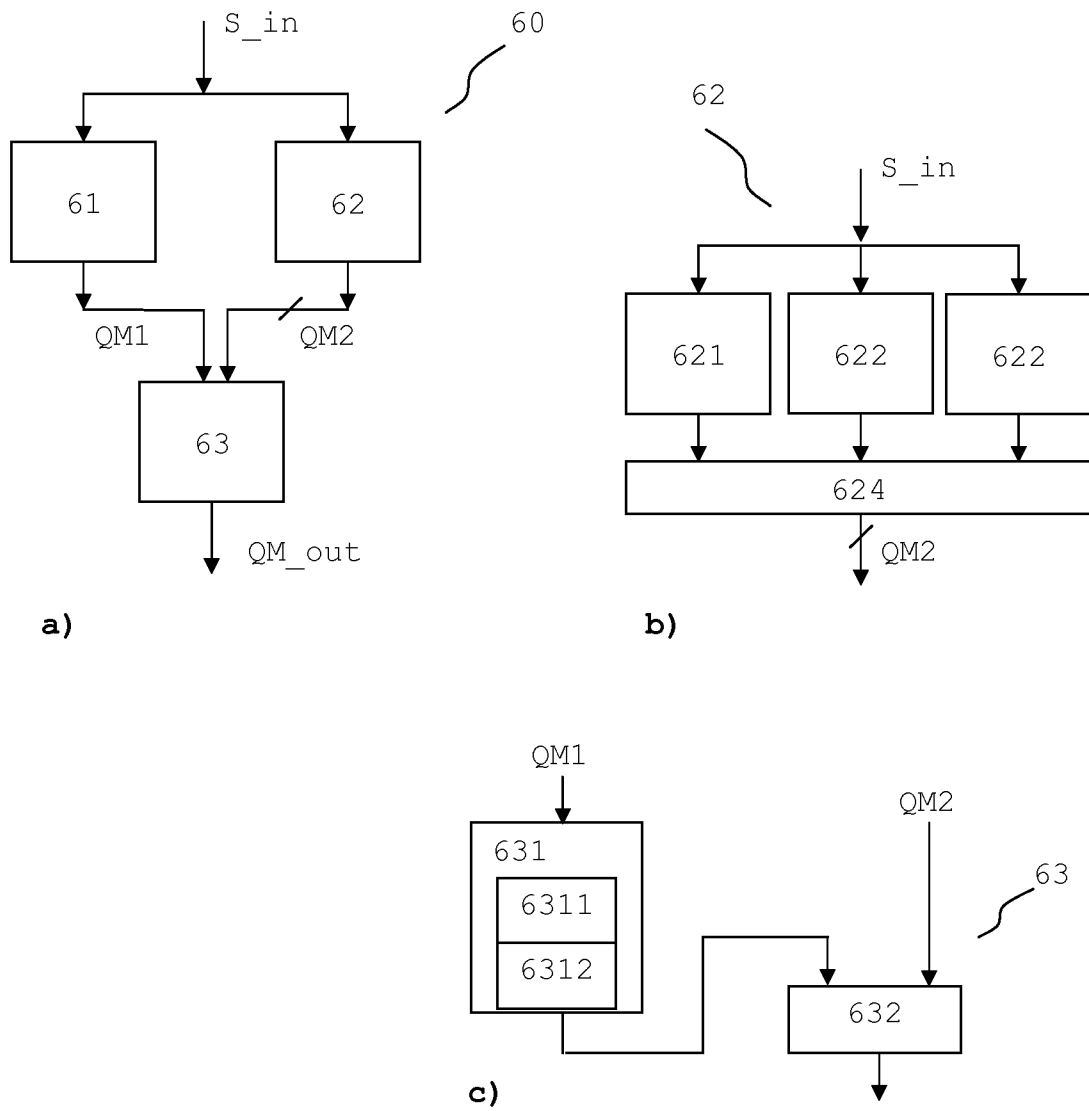


a)



b)

Fig.5

**Fig. 6**

INTERNATIONAL SEARCH REPORT

International application No
PCT/EP2010/053211

A. CLASSIFICATION OF SUBJECT MATTER INV. H04N7/26 H04N7/50 G06T5/00 G06T7/00 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 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.
X	ENGELKE U ET AL: "Perceptual-based Quality Metrics for Image and Video Services: A Survey" NEXT GENERATION INTERNET NETWORKS, 3RD EURONGI CONFERENCE ON, IEEE, PI LNKD-DOI:10.1109/NGI.2007.371215, 1 May 2007 (2007-05-01), pages 190-197, XP031176445 ISBN: 978-1-4244-0856-6	1,9,11
Y	cited in the application page 190, section I page 191, section II.B pages 194, sections III-III.C ----- -/--	2-8,10
☒ 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	
23 July 2010	02/08/2010	
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 André, Thomas	

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2010/053211

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	ICHIGAYA A ET AL: "A METHOD OF ESTIMATING CODING PSNR USING QUANTIZED DCT COEFFICIENTS" IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS FOR VIDEO TECHNOLOGY, IEEE SERVICE CENTER, PISCATAWAY, NJ, US LNKD- DOI:10.1109/TCSVT.2005.858745, vol. 16, no. 2, 1 February 2006 (2006-02-01), pages 251-259, XP001240857 ISSN: 1051-8215 cited in the application page 251, right-hand column, last paragraph pages 253-255, section III.A -----	2-4
Y	XIUHUA JIANG ET AL: "No-Reference Perceptual Video Quality Measurement for High Definition Videos Based on an Artificial Neural Network" COMPUTER AND ELECTRICAL ENGINEERING, 2008. ICCEE 2008. INTERNATIONAL CONFERENCE ON, IEEE, PISCATAWAY, NJ, USA, 20 December 2008 (2008-12-20), pages 424-427, XP031403397 ISBN: 978-0-7695-3504-3 cited in the application page 424, section I page 425, section II.B page 426, section III -----	5
Y	MARZILIANO P ET AL: "Perceptual blur and ringing metrics: application to JPEG2000" SIGNAL PROCESSING. IMAGE COMMUNICATION, ELSEVIER SCIENCE PUBLISHERS, AMSTERDAM, NL LNKD- DOI:10.1016/J.IMAGE.2003.08.003, vol. 19, no. 2, 1 February 2004 (2004-02-01), pages 163-172, XP004483133 ISSN: 0923-5965 pages 163-166, sections 1-2.1 -----	6-8
Y	FARIAS M C Q ET AL: "No-Reference Video Quality Metric Based on Artifact Measurements" IMAGE PROCESSING, 2005. ICIP 2005. IEEE INTERNATIONAL CONFERENCE ON GENOVA, ITALY 11-14 SEPT. 2005, PISCATAWAY, NJ, USA, IEEE, PISCATAWAY, NJ, USA LNKD- DOI:10.1109/ICIP.2005.1530348, vol. 3, 11 September 2005 (2005-09-11), pages 141-144, XP010851416 ISBN: 978-0-7803-9134-5 cited in the application the whole document ----- -/--	10

INTERNATIONAL SEARCH REPORT

International application No

PCT/EP2010/053211

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 2004/114685 A1 (KOULOHERIS JACK L [US] ET AL) 17 June 2004 (2004-06-17) figures 1-3 paragraph [0002] paragraphs [0011] - [0025] -----	1-11
A	LIGANG LU ET AL: "Full-reference video quality assessment considering structural distortion and no-reference quality evaluation of mpeg video" MULTIMEDIA AND EXPO, 2002. ICME '02. PROCEEDINGS. 2002 IEEE INTERNATIONAL CONFERENCE ON LAUSANNE, SWITZERLAND 26-29 AUG. 2002, PISCATAWAY, NJ, USA, IEEE, US, vol. 1, 26 August 2002 (2002-08-26), pages 61-64, XP010604306 ISBN: 978-0-7803-7304-4 page 61, section 1, last paragraph pages 62-63, section 3 figure 5 -----	1-11
A	WO 2007/066066 A2 (BRITISH TELECOMM [GB]; DAVIS ANDREW GORDON [GB]) 14 June 2007 (2007-06-14) page 1, line 3 - page 2, line 12 page 3, line 5 - page 10, line 14 -----	1-11
A	US 2007/217705 A1 (LEE SANGKEUN [US] ET AL) 20 September 2007 (2007-09-20) figures 4,5 paragraphs [0001] - [0005] paragraphs [0015] - [0035] -----	1-11

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No

PCT/EP2010/053211

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2004114685 A1	17-06-2004	NONE	
WO 2007066066 A2	14-06-2007	CA 2632403 A1 CN 101356827 A EP 1958454 A2 JP 2009518915 T US 2008317111 A1	14-06-2007 28-01-2009 20-08-2008 07-05-2009 25-12-2008
US 2007217705 A1	20-09-2007	NONE	