

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第5487455号
(P5487455)

(45) 発行日 平成26年5月7日(2014.5.7)

(24) 登録日 平成26年3月7日(2014.3.7)

(51) Int.Cl.

H04N 19/00 (2014.01)

F 1

H04N 7/13

Z

請求項の数 25 (全 29 頁)

(21) 出願番号 特願2011-539855 (P2011-539855)
 (86) (22) 出願日 平成20年12月23日 (2008.12.23)
 (65) 公表番号 特表2012-511843 (P2012-511843A)
 (43) 公表日 平成24年5月24日 (2012.5.24)
 (86) 國際出願番号 PCT/CA2008/002305
 (87) 國際公開番号 WO2010/066019
 (87) 國際公開日 平成22年6月17日 (2010.6.17)
 審査請求日 平成23年12月22日 (2011.12.22)
 (31) 優先権主張番号 12/333,406
 (32) 優先日 平成20年12月12日 (2008.12.12)
 (33) 優先権主張国 米国(US)

(73) 特許権者 509018915
 エコール・ドゥ・テクノロジー・スュペリュール
 カナダ国ケベック州 アーシュ3セー・1
 カ3 モントリオール、リュ・ノートルーダム・ウエスト 1100
 (74) 代理人 110000877
 龍華国際特許業務法人
 (72) 発明者 クーロン、ステファン
 カナダ国ケベック州 アーシュ3セー・1
 カ3 モントリオール、リュ・ノートルーダム・ウエスト 1100 エコール・ド
 ヴ・テクノロジー・スュペリュール内

最終頁に続く

(54) 【発明の名称】最適品質に近い品質での画像の複雑度の低いトランスコード方法およびシステム

(57) 【特許請求の範囲】

【請求項1】

端末での表示用に入力画像を出力画像にトランスコードする方法であって、
 (a 1) 前記入力画像の前記トランスコードの前に、トレーニング画像セットTをトランスコードすることによって、最大画像ファイルサイズ及び前記端末の解像度の組み合わせを有する前記端末の特徴を有するトランスコードパラメータを含むデータセットを生成する段階と、

前記データセットにおいて、前記トランスコードパラメータは、前記セット中のそれぞれのトレーニング画像と選択された前記端末の特徴について、選択された前記端末の特徴を満たしつつ、最高品質のトランスコードがなされたトレーニング画像を生成するように選択され、

(b 1) 前記入力画像の特徴を抽出する段階と、

(c 1) 前記段階(b 1)で抽出した前記入力画像の前記特徴と前記端末の特徴とを利用することで、前記段階(a 1)で生成された前記データセットから前記入力画像の前記トランスコードに対するトランスコードパラメータを選択する段階と、

(d 1) 前記段階(c 1)で選択された前記トランスコードパラメータを利用することで、前記入力画像を前記出力画像にトランスコードする段階と
を備える方法。

【請求項2】

前記段階(a 1)は、さらに

10

20

前記トレーニング画像の特徴に対応する前記トランスコードパラメータを含む前記データセットを生成する段階を有する請求項 1 に記載の方法。

【請求項 3】

前記段階 (b 1) は、

前記入力画像の符号化の忠実度を特徴付ける品質係数 Q F (I) 、画像のファイルサイズ S (I) 、画像の幅 W (I) 、および、画像の高さ H (I) を含む前記入力画像の前記特徴を抽出する段階を有する請求項 1 または 2 に記載の方法。

【請求項 4】

前記トランスコードパラメータを選択する前記段階 (c 1) は、

前記出力画像の符号化の忠実度を特徴付けるトランスコード品質係数 Q F T 、および、
前記トランスコードによる画像解像度の変化を特徴付けるトランスコードスケーリング係数 z T を選択する段階を有する請求項 1 から 3 のいずれか一項に記載の方法。
10

【請求項 5】

前記段階 (a 1) で生成される前記データセットは、N 個の多次元マトリックスを含み、前記 N 個の多次元マトリックスは、第 1 のマトリックス、第 2 のマトリックス、および第 3 のマトリックスのうち 1 以上を含み、

前記第 1 のマトリックスは、前記トレーニング画像セット T からの前記トレーニング画像の 1 以上のサブセットの最適品質係数の平均 Q F A を格納し、各 Q F A は前記トレーニング画像のそれぞれのサブセットが含む前記トレーニング画像のトランスコード中に利用される前記 Q F T の平均として決定され、前記最高品質を持つ前記トランスコードされたトレーニング画像が生成され、
20

前記第 2 のマトリックスは、前記トレーニング画像の 1 以上のサブセットの最適スケーリング係数の平均 z A を格納し、各 z A は前記トレーニング画像のそれぞれのサブセットが含む前記トレーニング画像のトランスコード中に利用される前記 z T の平均として決定され、前記最高品質を持つ前記トランスコードされたトレーニング画像が生成され、

前記第 3 のマトリックスは、前記トレーニング画像の 1 以上のサブセットの画質メトリックの平均 Q A を格納し、各 Q A は、前記トレーニング画像のそれぞれのサブセットについて、前記トレーニング画像と、それぞれ最高品質のトランスコードがなされたトレーニング画像との間の類似度を示す、品質メトリックの平均として決定される請求項 4 に記載の方法。
30

【請求項 6】

各マトリックスは、第 1 のインデックス、第 2 のインデックス、および第 3 のインデックスのうち 2 以上のインデックスからインデックスされ、

前記第 1 のインデックスは、前記入力画像の、前記入力画像の符号化の忠実度を特徴付ける品質係数 Q F (I) を示し、

前記第 2 のインデックスは、前記出力画像の視聴条件による画像の解像度の変化を特徴付ける視聴スケーリング係数 z V を示し、

前記第 3 のインデックスは、前記入力画像に対する前記出力画像の相対サイズを特徴付ける最大相対画像サイズ「 s m a x 」を示す請求項 5 に記載の方法。

【請求項 7】

前記段階 (a 1) は、

(a 7) 前記 Q F (I) 、前記 z V 、および前記「 s m a x 」の組み合わせを選択する段階と、

(a 7 - i) 前記 z V および前記「 s m a x 」の前記選択された組み合わせについて、品質係数が Q F (I) に等しい前記トレーニング画像のサブセットを利用して、前記最適品質係数の平均 Q F A および前記最適スケーリング係数の平均 z A を生成する段階と、

(a 7 - i i) 前記 Q F A を前記第 1 のマトリックスの、前記 z A を前記第 2 のマトリックスの、前記 Q F (I) 、前記 z V 、および前記「 s m a x 」の前記選択された組み合わせでインデックスされる位置にそれぞれ格納する段階と、

(a 7 - i i i) 前記 Q F (I) 、前記 z V 、および前記「 s m a x 」を、所定の間隔
40

50

で変化させて、前記 Q F (I)、前記 z V , および前記「 s m a x 」の全ての組み合わせが利用されるまで、前記段階 (a 7) を含む前記段階 (a 7) から前記段階 (a 7 - i i i) の各段階を繰り返す段階をさらに有する請求項 6 に記載の方法。

【請求項 8】

(a 8 - i) 前記 z V および前記「 s m a x 」の前記選択された組み合わせについて、品質係数が Q F (I) に等しい前記トレーニング画像のサブセットを利用して前記画質メトリックの平均 Q A を生成する段階であって、前記トレーニング画像の前記サブセットの各トレーニング画像についての前記最高品質のトランスクードがなされたトレーニング画像の前記品質メトリックを決定することを含む段階と、

(a 8 - i i) 前記 Q A を前記第 3 のマトリックスの、前記 Q F (I)、前記 z V , および前記「 s m a x 」の前記選択された組み合わせによりインデックスされる位置に格納する段階と、

前記段階 (a 8 - i) から前記段階 (a 8 - i i) を、前記段階 (a 7) の後であって前記段階 (a 7 - i i i) の前に実行する段階とを備える請求項 7 に記載の方法。 10

【請求項 9】

前記段階 (a 7 - i) は、

(a 9) 前記トレーニング画像の前記サブセットの各トレーニング画像について、前記最高品質のトランスクードがなされたトレーニング画像を生成する前記トランスクード品質係数 Q F T および前記トランスクードスケーリング係数 z T を決定する段階と、

(b 9) 前記段階 (a 9) で決定された前記トレーニング画像の前記トランスクード品質係数 Q F T および前記トランスクードスケーリング係数 z T とを利用して、前記最適品質係数の平均 Q F A および前記最適スケーリング係数の平均 z A を決定する段階とをさらに有する請求項 7 または請求項 8 に記載の方法。 20

【請求項 10】

前記段階 (a 9) は、

(a 10) 前記トランスクード品質係数 Q F T および前記トランスクードスケーリング係数 z T の組み合わせを含む前記トランスクードパラメータを選択する段階と、

(a 10 - i) 前記 Q F T および前記 z T の前記選択された組み合わせについて、前記トランスクードされたトレーニング画像を生成するべく前記トレーニング画像をトランスクードする段階と、

(a 10 - i i) 前記トレーニング画像と前記トランスクードされたトレーニング画像との間の類似度を示す、前記段階 (a 10 - i) で生成された前記トランスクードされたトレーニング画像の品質メトリックを決定する段階と、

(a 10 - i i i) 前記 Q F T および前記 z T を所定の間隔で変化させて、前記 Q F T および前記 z V の全ての組み合わせが利用されるまで、前記段階 (a 10) を含む前記段階 (a 10) から前記段階 (a 10 - i i i) の各段階を繰り返す段階と、

(a 10 - i v) 前記段階 (a 10 - i i i) で最高品質のメトリックを生成した前記 Q F T および前記 z T を選択する段階と

をさらに有する請求項 9 に記載の方法。 30

【請求項 11】

前記段階 (a 10 - i) は、

(a 11) 前記選択されたスケーリング係数 z T を適用することにより、前記トレーニング画像を伸張およびスケーリングする段階と、

(b 11) 前記選択された品質係数 Q F T を適用することにより、前記段階 (a 11) で生成された前記伸張およびスケーリングされた画像を圧縮する段階と

をさらに有する請求項 10 に記載の方法。 40

【請求項 12】

前記段階 (a 10 - i i) は、

(12 a) 画像 J' を生成するべく、リスケーリング係数 z R を適用することで前記トランスクードされたトレーニング画像を伸張およびスケーリングする段階と、

(12b) 画像 L' を生成するべく、前記視聴スケーリング係数 zV を適用することで前記トレーニング画像を伸張およびスケーリングする段階と、

(12c) 前記画像 L' および前記画像 J' を利用することで、前記トレーニング画像と前記トランスクードされたトレーニング画像との間の前記類似度を示す前記品質メトリックを生成する段階と

をさらに有する請求項10または11に記載の方法。

【請求項13】

前記 zR は zV / zT に等しく、

前記段階(12c)は、前記画像 J' および前記画像 L' を利用することで、Structural Similarity (SSIM) インデックスを決定する段階を有する請求項12に記載の方法。10

【請求項14】

前記段階(a8-i)は、

(a15) 前記第1のマトリックスに格納されている前記最適品質係数の平均 QFA と、前記第2のマトリックスに格納されている前記最適スケーリング係数の平均 zA を取得する段階と、

(a15-i) 前記トランスクードされたトレーニング画像を生成するべく、前記段階(a15)で取得された前記 QFA および前記 zA を利用して前記トレーニング画像をトランスクードする段階と、

(a15-ii) 前記トランスクードされたトレーニング画像が、前記選択された前記端末の特徴を満たすか否かを判断する段階と、20

(a15-iii) 前記トランスクードされたトレーニング画像が前記選択された前記端末の特徴を満たすまで、前記第1および前記第2のマトリックスで利用可能な前記「 $smax$ 」の次に小さい値で、前記段階(a15)を含む前記段階(a15)から前記段階(a15-iii)の各段階を繰り返す段階と、

(b15) 前記選択された前記端末の特徴を満たす段階(a15-i)で生成された前記トランスクードされたトレーニング画像の前記品質メトリックを決定する段階と

をさらに有する請求項8から13のいずれか一項に記載の方法。

【請求項15】

前記段階(d1)は、30

(a17) 前記端末の前記特徴、および、前記段階(b1)で抽出された $H(I)$ 、 $W(I)$ 、 $S(I)$ 、 $QF(I)$ 、前記入力画像の前記特徴を利用して、前記 zV および前記「 $smax$ 」を決定する段階と、

(b17) 前記段階(b1)で取得された前記 $QF(I)$ と前記段階(a17)により決定された前記 zV および前記「 $smax$ 」とによりインデックスされた前記第2のマトリックスに格納されている最適スケーリング係数の平均 zA と、前記第1のマトリックスに格納されている前記最適品質係数の平均 QFA を取得する段階と、

(b17-i) 前記出力画像を生成するべく、前記トランスクード品質係数 QFT および前記トランスクードスケーリング係数 zT として前記段階(b17)で取得された前記 QFA および前記 zA を利用して前記入力画像をトランスクードする段階と、40

(b17-ii) 前記出力画像が前記端末の前記特徴を満たすか否かをチェックする段階と、

(b17-iii) 前記出力画像が前記端末の前記特徴を満たすまで、前記第1および前記第2のマトリックスで利用可能な前記「 $smax$ 」の次に小さい値で、前記段階(b17)を含む前記段階(b17)から前記段階(b17-iii)の各段階を繰り返す段階と、

(c17) 前記第3のマトリックスの前記 $QF(I)$ がインデックスする位置に格納されている前記画質の平均 QA を取得する段階と

をさらに有し、

前記「 $smax$ 」および前記 zV は、前記端末の前記特徴を満たす前記段階(b17-50

i) で生成された前記出力画像に対応している
請求項 6 から 14 のいずれか一項に記載の方法。

【請求項 16】

前記端末の特徴は、最大画像ファイルサイズ S (D) 、前記端末の画面の幅 W (D) 、
および前記端末の画面の高さ H (D) を含み、前記方法は、

(a 18) 前記 z V を、(W (D) / W (I)) , (H (D) / H (I)) および 1 のうちの最小値に設定する段階と、

(b 18) 前記「 s m a x 」を、(S (D) / S (I)) および 1 のうちの最小値に設定する段階とをさらに備える請求項 6 から 15 のいずれか一項に記載の方法。

【請求項 17】

端末での表示用に入力画像を出力画像にトランスコードする、プロセッサおよびコンピュータ可読格納媒体を備えるシステムであって、前記システムは、

(a 24) トレーニング画像を含むトレーニング画像セット T を有する、前記コンピュータ可読格納媒体に格納されるトレーニング画像レポジトリと、

(b 24) 前記コンピュータ可読格納媒体に格納されたコンピュータ可読命令を有するオフラインデータセット生成器であって、前記オフラインデータセット生成器は、前記入力画像の前記トランスコードの前に、トレーニング画像の前記トレーニング画像セット T をトランスコードすることによって、最大画像ファイルサイズ及び前記端末の解像度の組み合わせを有する前記端末の特徴を有するトランスコードパラメータを含むデータセットを生成し、

前記データセットにおいて、前記トランスコードパラメータは、前記トレーニング画像セット T 中のそれぞれのトレーニング画像と選択された前記端末の特徴について、選択された前記端末の特徴を満たしつつ、最高品質のトランスコードがなされたトレーニング画像を生成するように選択され、

(c 24) 前記オフラインデータセット生成器 (b 24) が生成する前記データセットを格納する、コンピュータ可読格納媒体に格納されるデータセットレポジトリと、

(d 24) 前記入力画像の特徴を抽出し、前記端末の特徴と前記入力画像の前記特徴と を利用して前記データセットレポジトリ (c 24) に格納されている前記データセットから前記入力画像の前記トランスコードに対するトランスコードパラメータを選択し、前記選択されたトランスコードパラメータを利用して前記入力画像を前記出力画像にトランスコードするように構成する コンピュータ可読命令をコンピュータ可読格納媒体に有するオンライントランスコーダと

を備えるシステム。

【請求項 18】

前記オンライントランスコーダは、さらに前記入力画像の前記特徴に対応する前記トランスコードパラメータを選択するように構成され、

前記トランスコードパラメータは、さらに前記出力画像の符号化の忠実度を特徴付けるトランスコード品質係数 Q F T と、前記トランスコードによる画像の解像度の変化を特徴付けるトランスコードスケーリング係数 z T とを含む請求項 17 に記載のシステム。

【請求項 19】

前記データセットは N 個の多次元マトリックスを含み、前記 N 個の多次元マトリックスは、第 1 のマトリックス、第 2 のマトリックス、および第 3 のマトリックスのうち一以上を含み、

前記第 1 のマトリックスは、前記トレーニング画像セット T からの前記トレーニング画像の 1 以上のサブセットの最適品質係数の平均 Q F A を格納し、各 Q F A は前記トレーニング画像のそれぞれのサブセットが含むトレーニング画像のトランスコード中に利用される前記 Q F T の平均として決定され、前記最高品質を持つ前記トランスコードされたトレーニング画像が生成され、

前記第 2 のマトリックスは、前記トレーニング画像の 1 以上のサブセットの最適スケーリング係数の平均 z A を格納し、各 z A は前記トレーニング画像のそれぞれのサブセット

10

20

30

40

50

が含むトレーニング画像のトランスクード中に利用される前記 z_T の平均として決定され、前記最高品質を持つ前記トランスクードされたトレーニング画像が生成され、

前記第3のマトリックスは、前記トレーニング画像の1以上のサブセットの画質メトリックの平均 Q_A を格納し、各 Q_A は、トレーニング画像のそれぞれのサブセットについて、前記トレーニング画像と、それぞれ対応する最高品質のトランスクードがなされたトレーニング画像との間の類似度を示す、品質メトリックの平均として決定される請求項18に記載のシステム。

【請求項20】

各マトリックスは、第1のインデックス、第2のインデックス、および第3のインデックスのうち2以上のインデックスからインデックスされ、

前記第1のインデックスは、前記入力画像の符号化の忠実度を特徴付ける品質係数 Q_F (I) を示し、

前記第2のインデックスは、前記出力画像の視聴条件による画像の解像度の変化を特徴付ける視聴スケーリング係数 z_V を示し、

前記第3のインデックスは、前記入力画像に対する前記出力画像の相対サイズを特徴付ける最大相対画像サイズ「 s_{max} 」を示す請求項19に記載のシステム。

【請求項21】

前記オフラインデータセット生成器 (b24) は、

(a29) 前記トレーニング画像セットTのトレーニング画像の特徴を抽出する画像特徴抽出ユニットと、

(b29) 前記 $Q_F T$ と前記 z_T とを選択し、前記画像特徴抽出ユニット (a24) が抽出した前記特徴を利用して前記最適品質係数の平均 $Q_F A$ および前記最適スケーリング係数の平均 z_A を決定する第1の品質認識パラメータ選択ユニットと、

(c29) 前記第1の品質認識パラメータ選択ユニット (b29) が選択した前記 $Q_F T$ および前記 z_T を利用して、前記トレーニング画像を、トランスクードされたトレーニング画像にトランスクードするトランスクードユニットと、

(d29) 前記トランスクードユニット (c29) が生成した前記トランスクードされたトレーニング画像の、前記トレーニング画像と前記トランスクードされたトレーニング画像との間の類似度を示す前記品質メトリックを決定する品質評価ユニットとを有する請求項19または20に記載のシステム。

【請求項22】

前記トランスクードユニット (c29) は、

(a31) 前記スケーリング係数 z_T を前記トレーニング画像に適用することで、伸張およびスケーリングされたトレーニング画像を生成する伸張およびスケーリングユニットと、

(b31) 前記トランスクード品質係数 $Q_F T$ を前記伸張およびスケーリングされたトレーニング画像に適用することで、前記トランスクードされたトレーニング画像を生成する圧縮ユニットとを有する請求項21に記載のシステム。

【請求項23】

前記品質評価ユニット (d29) は、

(a32) リスキーリング係数 z_R を適用することで、前記トランスクードされたトレーニング画像を伸張およびスケーリングする第2の伸張およびスケーリングユニットと、

(b32) 視聴スケーリング係数 z_V を適用することで、前記トレーニング画像を伸張およびスケーリングする第3の伸張およびスケーリングユニットと、

(c32) 前記トランスクードされたトレーニング画像の前記品質メトリックを計算する品質メトリック計算ユニットとを有し、前記トレーニング画像と前記トランスクードされたトレーニング画像との間の類似度を示す請求項21または22に記載のシステム。

【請求項24】

前記品質メトリック計算ユニット (c32) は、

前記トレーニング画像と前記トランスクードされたトレーニング画像との間の類似性を

10

20

30

40

50

示す Structural Similarity (SSIM) インデックスを決定する計算手段を有する請求項 23 に記載のシステム。

【請求項 25】

コンピュータに、請求項 1 から 16 のいずれか一項に記載された前記方法の前記段階それぞれを実行させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像トランスコード分野に係り、特に、最適品質に近い品質での画像の複雑度の低いトランスコード方法およびシステムに係る。

10

【背景技術】

【0002】

テキスト、オーディオ、およびビデオを含む様々なコンテンツの種類のマルチメディアが、ネットワークオペレータに対して顕著なビジネスおよび収入のチャンスを提供している。より高い帯域幅の利用可能性およびパケット交換インターネットプロトコル (IP) 技術の利用により、テキスト、音声、静止およびアニメーショングラフィック、写真、ビデオクリップ、および音楽の様々な組み合わせが含まれる、より内容豊富なコンテンツの送信が可能となっている。この市場から利益を得るには、この市場の潜在的なネットワークオペレータが、品質および信頼性に関する顧客の期待に応える必要がある。移動端末、携帯電話、コンピュータ、その他の電子デバイスを含む今日の異機種環境にあるネットワーク間でマルチメディアアプリケーションを利用するには、サーバレベルでのメディアのトランスコードが重要となる。個々のデバイスではリソースに制約があることが多く、メディア自身を適合する機能を持つことが稀なために、サービスプロバイダレベルでメディアの適合およびトランスコードを行う必要がある。今後数年のトラフィックの成長は新たなハードウェアのみでそのスピードについてゆくには著しすぎるために、これはサービスプロバイダにとって重要な問題となる。

20

【0003】

例えばマルチメディアメッセージサービスは、サーバ側の適合を提供して端末間の互換性を保証する。マルチメディアメッセージに必要な適合については、S. Coulombe および G. Grassel による「Multimedia adaptation for the multimedia messaging service」、I E E E 通信雑誌、第 42 巻、7 号、120 - 126 ページ、2004 年 7 月が説明している。もっとも頻繁に生じる画像関連の互換性の問題は、画像フォーマットに関するものではなく、トラフィックの大半が J P E G および G I F 画像に関するために、解像度またはファイルサイズが受信端末の機能を超える、というものである。例えば、一部の移動電話ではメモリが限られていることから、画像を一定のサイズまたは解像度に収まるようにして、受信および表示を可能にする必要がある。さらに、画像の高さおよび幅を、決められた特徴セットを有する端末が収容可能なものにする必要がある。画像の解像度、またはスケーリングを変更して、端末の機能に整合させる方法としては、公知の解決法が存在する。しかし、画質を端末の制約に対して最適化することは、J P E G で利用されている品質係数 (Q F) (または G I F で利用されている色数)、知覚品質、および圧縮ファイルサイズの間に、確立された関連性がないことから、未だ解決されていない課題である。単なる解像度の適合ではなく、ファイルサイズを低減させるための追加手段としてスケーリングを利用しようとすると、この問題はさらに大きな課題となる。ビジュアルコンテンツのファイルサイズ (またはビットレート) の問題を調べた研究は既に幾つか存在している。

30

【0004】

数例としては Z. Lei および N.D. Georganas による「Accurate bit allocation and rate control for DCT domain video transcoding」なる名称の研究論文、IEEE CCECE 予稿集 2002 年、Canadian Conference on Electrical and Computer Engineering, 2002 年、第 2 巻、ページ 968-973、「Signal Processing: Image Communication」第 18 巻、

40

50

8号、ページ621-639、2003年9月内で発表されたJ. Ridgeによる「Efficient transform-domain size and resolution reduction of images」なる名称の研究論文、および、V. Ratnakar and V. Ivashinによる「File size bounded JPEG transcoder」なる名称の米国特許第6,233,359号明細書、2001年5月が含まれる。これら研究の結果は、スケーリングではなくて、量子化パラメータを適合させることによって低減を達成することができるこことを示している。殆どの研究では、解像度が幾つかの予め定義されたフォーマットに限られることの多い低ビットレートのビデオのコンテクストからの話なので、このことは納得がいく。しかし、静止画像符号化のコンテクストであっても、スケーリングが適合のための戦略としては考えられていない。例えば、Ridgeは、スケーリングしてからJPEG画像のファイルサイズを低減させる優れた方法を提供しているが、スケーリングと品質の低減とを組み合わせた評価を行っていない。ユーザ体験を最大化するための最良の戦略というものは、画像を低いQFで単に再度圧縮するのではなくて、より高いQFで画像を縮尺して圧縮する必要があることから、これは大きな欠点であると思われる。

【0005】

Steven PigeonおよびStephane Coulombeによる「Computationally efficient algorithms for predicting the file size of JPEG images subject to changes of quality factor and scaling」、24th Queen's Biennial Symposium on Communications, Queen's University, Kingston, Canada, 2008では、スケーリングおよびQF変化を受けるJPEG画像の圧縮ファイルサイズを推定する方法が提示されている。QFおよびスケーリングの幾らかの組み合わせによって、同様の近似したファイルサイズが生じ、どの組み合わせがユーザ体験または画質を最大化するか、という問題が生じる。本発明で前述した課題がいかに解決されるかを述べる前に、通常の先行技術によるトランスコード環境を簡単に説明する。JPEGは、MMSメッセージに含まれる画像圧縮技術として一般に普及している。JPEG規格により、通常利用されている画像圧縮方法が提供される。よく知られているように、JPEG圧縮は「損失が多い」、つまり、圧縮された画像が原画像に含まれていたデジタル情報を100%含まないことがある。情報の損失は、圧縮中に品質係数を設定することにより制御可能である。品質係数が低いことは、圧縮度が高いことに等しく、一般的には画像のファイルサイズが小さいことを意味する。逆に、品質係数が高いことは、画像のファイルサイズが大きいことを意味しており、一般的には、知覚される画像の「品質」が高いことを意味する。

【0006】

例えば、図1は、発信元ノード102、サービス配信プラットフォーム104、発信先ノード106、および適合エンジン108を含むマルチメディアメッセージ(MMS)システムアーキテクチャ100を示している。発信元ノード102は、ネットワーク「A」110を介してサービス配信プラットフォーム104と通信することができる。同様にして、発信先ノード106は、ネットワーク「B」112を介してサービス配信プラットフォーム104と通信することができる。ネットワーク「A」および「B」は、可能な接続セットを示す例示にすぎず、多くの他の構成も可能である。例えば発信元および発信先ノード(102および106)が、単一のネットワークを介してサービス配信プラットフォーム104と通信してもよく、発信元ノード102が仲介ネットワークを介さずにサービス配信プラットフォーム104と直接接続されていてもよい。

【0007】

適合エンジン108は、サービス配信プラットフォーム104と、図1に示すようにリンク114を介して直接接続されても、あるいは、ネットワーク経由で接続されていても、またはあるいは、サービス配信プラットフォーム104に埋め込まれていてもよい。

【0008】

単純なケースとして、発信元ノード102が、発信先ノード106宛ての(マルチメディア)メッセージを送信する場合を考える。メッセージはネットワーク「A」110を介してサービス配信プラットフォーム104に転送され、そこからメッセージは、ネットワーク「B」112を介して発信先ノード106へと送られる。発信元および発信先ノード

(102および106)は、例えば無線デバイスであってよく、ネットワーク「A」および「B」(110および112)は、この場合に無線ネットワークであり、サービス配信プラットフォーム104はマルチメディアメッセージ転送サービスを提供してよい。

【0009】

別の例として、発信元ノード102が、データネットワークを介してサービス配信プラットフォーム104に接続されたコンテンツプロバイダのサーバである場合を考える。この場合、ネットワーク「A」110がインターネットであり、ネットワーク「B」112が、無線デバイスであってよい発信先ノード106に機能を提供する無線ネットワークであってよい。

【0010】

マルチメディアメッセージサービス(MMS)のサーバ側の適合の概略は、Stephane C oulombeおよびG. Grasselによる「Multimedia Adaptation for the Multimedia Messaging Service」、IEEE通信雑誌、第42巻、7号、120 - 126ページ、2004年7月に述べられているので、これを参照されたい。

【0011】

特に画像の場合、発信元ノード102が送信するメッセージには、画像(特にJPEG符号化画像)が含まれてよい。ピクセル数による画像の高さまたは幅(これらはつまり画像の解像度である)が発信先ノード106の表示デバイスまたは端末のサイズまたは解像度を超える等の理由から、発信先ノード106の機能に元のフォームの画像表示機能が含まれていない場合がある。発信先ノード106がこの画像を受信して表示するためには、画像は、発信先ノード106に配信される前に、適合エンジン108の画像トランスコード116により修正される必要がある。画像トランスコード116による画像の修正は、通常、スケーリング(つまり画像の改造度の変更)および圧縮を含んでよい。

【0012】

従って、本産業では、前述した先行技術の制約を克服でき、画質およびトランスコード速度も考慮に入れた、画像のトランスコード方法およびシステムの向上が望まれている。

【発明の概要】

【0013】

ゆえに、本発明の目的の一つは、最適品質に近い品質での画像の複雑度の低いトランスコード(例えば、MMSメッセージに含まれている画像のトランスコード)方法およびシステムを提供することである。

【0014】

端末での表示用に入力画像を出力画像にトランスコードする方法であって、(a1)トランスコードの前に、各トレーニング画像について最高品質のトランスコードされたトレーニング画像を生成するべくトレーニング画像セットT内のトレーニング画像のセットの最適品質のトランスコードを行うことで生成されるトランスコードパラメータを含み、且つ、端末の特徴を満たすデータセットを生成する段階と、(b1)入力画像の特徴を抽出して、端末の特徴を取得する段階と、(c1)段階(b1)で抽出した入力画像の特徴と端末の特徴とを利用することで、段階(a1)で生成されたデータセットからトランスコードパラメータを選択する段階と、(d1)段階(c1)で選択されたトランスコードパラメータを利用して、入力画像を出力画像にトランスコードする段階とを備える方法を提供する。段階(a1)は、端末の特徴と入力画像の特徴との組み合わせに対応するトランスコードパラメータを含むデータセットを生成する段階を有する。段階(b1)は、入力画像の符号化の忠実度を特徴付ける品質係数QFT(I)、画像のファイルサイズS(I)、画像の幅W(I)、および、画像の高さH(I)を含む入力画像の特徴を抽出する段階を有する。

【0015】

トランスコードパラメータを選択する段階(c1)は、出力画像の符号化の忠実度を特徴付けるトランスコード品質係数QFT、および、トランスコードによる画像解像度の変化を特徴付けるトランスコードスケーリング係数zTを選択する段階を有する。段階(a

10

20

30

40

50

1) で生成されるデータセットは、N 個の多次元マトリックスを含み、第1のマトリックス、第2のマトリックス、および第3のマトリックスのうち1以上を含み、第1のマトリックスは、トレーニング画像セットTからのトレーニング画像の1以上のサブセットの最適品質係数の平均QFTを格納し、各QFTはトレーニング画像のそれぞれのサブセットが含むトレーニング画像の最適品質のトランスクード中に利用されるQFTの平均として決定され、第2のマトリックスは、トレーニング画像の1以上のサブセットの最適スケーリング係数の平均zAを格納し、各zAはトレーニング画像のそれぞれのサブセットが含むトレーニング画像の最適品質のトランスクード中に利用されるzTの平均として決定され、第3のマトリックスは、トレーニング画像の1以上のサブセットの最適画質メトリックの平均QAを格納し、各QAは、トレーニング画像のそれぞれのサブセットについて、トレーニング画像と、最適品質のトランスクード中に生成されるそれぞれ最高品質のトランスクードされたトレーニング画像との間の類似度を示す、品質メトリックの平均として決定される。各マトリックスは、第1のインデックス、第2のインデックス、および第3のインデックスのうち2以上のインデックスからインデックスされ、第1のインデックスは、入力画像の品質係数QF(I)を示し、第2のインデックスは、出力画像の視聴条件による画像の解像度の変化を特徴付ける視聴スケーリング係数zVを示し、第3のインデックスは、入力画像に対する出力画像の相対サイズを特徴付ける最大相対画像サイズ「smax」を示す。10

【0016】

段階(a1)は、(a7)QF(I)、zV、および「smax」の組み合わせを選択する段階と、(a7-i)zVおよび「smax」の選択された組み合わせについて、品質係数がQF(I)に等しいトレーニング画像のサブセットを利用して、最適品質係数の平均QFAおよび最適スケーリング係数の平均zAを生成する段階と、(a7-ii)QFAを第1のマトリックスの、zAを第2のマトリックスの、QF(I)、zV、および「smax」の選択された組み合わせでインデックスされる位置にそれぞれ格納する段階と、(a7-iii)QF(I)、zV、および「smax」を、所定の間隔で変化させて、QF(I)、zV、および「smax」の全ての組み合わせが利用されるまで、段階(a7)を含む段階(a7)から段階(a7-iii)の各段階を繰り返す段階をさらに有する。方法は、(a8-i)zVおよび「smax」の選択された組み合わせについて、品質係数がQF(I)に等しいトレーニング画像のサブセットを利用して画質メトリックの平均QAを生成する段階であって、トレーニング画像のサブセットの各トレーニング画像についての最高品質のトランスクードされたトレーニング画像の品質メトリックを決定することを含む段階と、(a8-ii)QAを第3のマトリックスの、QF(I)、zV、および「smax」の選択された組み合わせによりインデックスされる位置に格納する段階と、段階(a8-i)から段階(a8-ii)を、段階(a7)の後であって段階(a7-iii)の前に実行する段階とを備える。20

【0017】

段階(a7-i)は、(a9)トレーニング画像のサブセットの各トレーニング画像について、最高品質のトランスクードされたトレーニング画像を生成するトランスクード品質係数QFTおよびトランスクードスケーリング係数zTを決定する段階と、(b9)段階(a9)で決定されたトレーニング画像のトランスクード品質係数QFTおよびトランスクードスケーリング係数zTとを利用して、最適品質係数の平均QFAおよび最適スケーリング係数の平均zAを決定する段階とをさらに有する。段階(a9)は、(a10)トランスクード品質係数QFTおよびトランスクードスケーリング係数zTの組み合わせを含むトランスクードパラメータを選択する段階と、(a10-i)QFTおよびzTの選択された組み合わせについて、トランスクードされたトレーニング画像を生成するべくトレーニング画像をトランスクードする段階と、(a10-ii)トレーニング画像とトランスクードされたトレーニング画像との間の類似度を示す、段階(a10-i)で生成されたトランスクードされたトレーニング画像の品質メトリックを決定する段階と、(a10-iii)QFTおよびzTを所定の間隔で変化させて、QFTおよびzVの全ての304050

組み合わせが利用されるまで、段階 (a 1 0) を含む段階 (a 1 0) から段階 (a 1 0 - i i i) の各段階を繰り返す段階と、(a 1 0 - i v) 段階 (a 1 0 - i i i) で最高品質のメトリックを生成した Q F T および z T を選択する段階とをさらに有する。段階 (a 1 0 - i) は、(a 1 1) 選択されたスケーリング係数 z T を適用することにより、トレーニング画像を伸張およびスケーリングする段階と、(b 1 1) 選択された品質係数 Q F T を適用することにより、段階 (a 1 1) で生成された伸張およびスケーリングされた画像を圧縮する段階とをさらに有する。段階 (a 1 0 - i i) は、(1 2 a) 画像 J' を生成するべく、リスケーリング係数 z R を適用することでトランスクードされたトレーニング画像を伸張およびスケーリングする段階と、(1 2 b) 画像 L' を生成するべく、視聴スケーリング係数 z V を適用することでトレーニング画像を伸張およびスケーリングする段階と、(1 2 c) 画像 L' および画像 J' を利用することで、トレーニング画像とトランスクードされたトレーニング画像との間の類似度を示す品質メトリックを生成する段階とをさらに有する。

【0 0 1 8】

z R は z V / z T に等しい。段階 1 2 (c) は、画像 J' および画像 L' を利用することで、構造的な S I M i l a r i t y (S S I M) インデックスを決定する段階を有する。段階 (a 8 - i) は、(a 1 5) 第 1 のマトリックスに格納されている最適品質係数の平均 Q F A と、第 2 のマトリックスに格納されている最適スケーリング係数の平均 z A を取得する段階と、(a 1 5 - i) トランスクードされたトレーニング画像を生成するべく、段階 (a 1 5) で取得された Q F A および z A を利用してトレーニング画像をトランスクードする段階と、(a 1 5 - i i) トランスクードされたトレーニング画像が、端末の特徴を満たすか否かを判断する段階と、(a 1 5 - i i i) トランスクードされたトレーニング画像が端末の特徴を満たすまで、第 1 および第 2 のマトリックスで利用可能な「 s m a x 」の次に小さい値で、段階 (a 1 5) を含む段階 (a 1 5) から段階 (a 1 5 - i i i) の各段階を繰り返す段階と、(b 1 5) 端末の特徴を満たす段階 (a 1 5 - i) で生成されたトランスクードされたトレーニング画像の品質メトリックを決定する段階とをさらに有する。段階 (a 1 5 - i i) は、トランスクードされたトレーニング画像の画像ファイルサイズが、端末がサポートする最大画像ファイルサイズ S (D) 以下であるかを判断する段階を有する。

【0 0 1 9】

段階 (d 1) は、(a 1 7) 端末の特徴、および、段階 (b 1) で抽出された H (I) 、 W (I) 、 S (I) 、 Q F (I) 、入力画像の特徴を利用して、z V および「 s m a x 」を決定する段階と、(b 1 7) 段階 (b 1) で取得された Q F (I) と段階 (a 1 7) により決定された z V および「 s m a x 」とによりインデックスされた第 2 のマトリックスに格納されている最適スケーリング係数の平均 z A と、第 1 のマトリックスに格納されている最適品質係数の平均 Q F A を取得する段階と、(b 1 7 - i) 出力画像を生成するべく、トランスクード品質係数 Q F T およびトランスクードスケーリング係数 z T として段階 (b 1 7) で取得された Q F A および z A を利用して入力画像をトランスクードする段階と、(b 1 7 - i i) 出力画像が端末の特徴を満たすか否かをチェックする段階と、(b 1 7 - i i i) 出力画像が端末の特徴を満たすまで、第 1 および第 2 のマトリックスで利用可能な「 s m a x 」の次に小さい値で、段階 (b 1 7) を含む段階 (b 1 7) から段階 (b 1 7 - i i i) の各段階を繰り返す段階と、(c 1 7) 第 3 のマトリックスの Q F (I) がインデックスする位置に格納されている画質の平均 Q A を取得する段階とをさらに有し、「 s m a x 」および z V は、端末の特徴を満たす段階 (b 1 7 - i) で生成された出力画像に対応している。

【0 0 2 0】

端末の特徴は、最大画像ファイルサイズ S (D) 、端末の幅 W (D) 、および端末の高さ H (D) を含み、段階 (a 1 7) はさらに、出力画像の符号化の忠実度を特徴付けるトランスクード品質係数 Q F T および画像の変化を特徴付けるトランスクードスケーリング係数 z T を選択する段階を有し、(a 1 8) z V を、(W (D) / W (I)) , H (D)

10

20

30

40

50

/ H (I)) および 1 のうちの最小値に設定する段階と、(b 1 8) 「 s m a x 」を、(S (D) / S (I)) および 1 のうちの最小値に設定する段階とを有する。段階 (a 1) は、(a 1 9) トランスコード品質係数 Q F T およびトランスコードスケーリング係数 z T の組み合わせを含むトランスコードパラメータを選択する段階と、(a 1 9 - i) 選択された Q F T および z T の組み合わせについて、トレーニング画像をトランスコードして、トランスコードされたトレーニング画像を生成する段階と、(a 1 9 - i i) 段階 (a 1 9 - i) で生成されたトランスコードされたトレーニング画像の、トレーニング画像とトランスコードされたトレーニング画像との間の類似度を示す品質メトリックを決定する段階と、(a 1 9 - i i i) Q F T および z T を所定の間隔で変化させて、Q F T および z T の全ての組み合わせが利用されるまで、(a 1 9) を含む (a 1 9) から (a 1 9 - i i) の各段階を繰り返す段階と、(a 1 9 - i v) 段階 (a 1 9 - i i) で最高品質のメトリックを生成した Q F T および z T を選択する段階とを有する。
10

【 0 0 2 1 】

段階 a - 1 9 (i) はさらに、段階 (a 2 0) 選択されたスケーリング係数 z T を適用してトレーニング画像を伸張およびスケーリングする段階と、(b 2 0) 選択された品質係数 Q F T を適用することで、段階 (a 2 0) で生成された伸張およびスケーリングされたトレーニング画像を圧縮する段階とを有する。段階 (a 1 9 - i i) はさらに、(a 2 1) リスクーリング係数 z R を適用することで、トランスコードされたトレーニング画像を伸張およびスケーリングして画像 J' を生成する段階と、(b 2 1) 出力画像の視聴条件による画像の解像度の変化を特徴付ける視聴スケーリング係数 z V を適用することで、トランスコードされたトレーニング画像を伸張およびスケーリングして画像 L' を生成する段階と、(c 2 1) 画像 J' および画像 L' をそれぞれ利用して、トレーニング画像とトランスコードされたトレーニング画像の間の類似度を示す品質メトリックを生成する段階とを有する。z R は z V / z T に等しい。段階 (c 2 1) は、画像 J' および画像 L' を利用することで、構造的な S I M i l a r i t y (S S I M) インデックスを決定する段階を有する。
20

【 0 0 2 2 】

端末での表示用に入力画像を出力画像にトランスコードする、プロセッサおよびコンピュータ可読格納媒体を備えるシステムであって、システムは、(a 2 4) トレーニング画像を含むトレーニング画像セット T を有する、コンピュータ可読格納媒体に格納されるトレーニング画像レポジトリと、(b 2 4) コンピュータ可読格納媒体に格納されたコンピュータ可読命令を有するオンラインデータセット生成器であって、入力画像のトランスコードの前に各トレーニング画像について最高品質のトランスコードされたトレーニング画像を生成するべくトレーニング画像セット T 内のトレーニング画像のセットの最適品質のトランスコードを行うことで生成されるトランスコードパラメータを含み、且つ、端末の特徴を満たすデータセットを生成するオンラインデータセット生成器と、(c 2 4) オフラインデータセット生成器 (b 2 4) が生成するデータセットを格納する、コンピュータ可読格納媒体に格納されるデータセットレポジトリと、(d 2 4) 端末の特徴と入力画像の特徴とを利用してデータセットレポジトリ (c 2 4) に格納されているデータセットからトランスコードパラメータを選択して、選択されたトランスコードパラメータを利用して入力画像を出力画像にトランスコードするコンピュータ可読命令をコンピュータ可読格納媒体に有するオンライントランスコーダとを備えるシステムが提供される。
30
40

【 0 0 2 3 】

データセットは、端末の特徴と入力画像の特徴との組み合わせに対応するトランスコードパラメータを含む。トランスコードパラメータは、出力画像の符号化の忠実度を特徴付けるトランスコード品質係数 Q F T と、トランスコードによる画像の解像度の変化を特徴付けるトランスコードスケーリング係数 z T とを含む。データセットは N 個の多次元マトリックスを含み、N 個の多次元マトリックスは、第 1 のマトリックス、第 2 のマトリックス、および第 3 のマトリックスのうち一以上を含み、第 1 のマトリックスは、トレーニング画像セット T からのトレーニング画像の 1 以上のサブセットの最適品質係数の平均 Q F
50

Aを格納し、各QFAはトレーニング画像のそれぞれのサブセットが含むトレーニング画像の最適品質のトランスクード中に利用されるQFTの平均として決定され、第2のマトリックスは、トレーニング画像の1以上のサブセットの最適なスケーリング係数の平均zAを格納し、各zAはトレーニング画像のそれぞれのサブセットが含むトレーニング画像の最適品質のトランスクード中に利用されるzTの平均として決定され、第3のマトリックスは、トレーニング画像の1以上のサブセットの最適画質メトリックの平均QAを格納し、各QAは、トレーニング画像のそれぞれのサブセットについて、トレーニング画像と、最適品質のトランスクード中に生成されるそれぞれ対応する最高品質のトランスクードされたトレーニング画像との間の類似度を示す、品質メトリックの平均として決定される。

10

【0024】

各マトリックスは、第1のインデックス、第2のインデックス、および第3のインデックスのうち2以上のインデックスからインデックスされ、第1のインデックスは、入力画像の符号化の忠実度を特徴付ける品質係数QF(I)を示し、第2のインデックスは、出力画像の視聴条件による画像の解像度の変化を特徴付ける視聴スケーリング係数zVを示し、第3のインデックスは、入力画像に対する出力画像の相対サイズを特徴付ける最大相対画像サイズ「 s_{max} 」を示す。オンラインデータセット生成器(b24)は、(a29)トレーニング画像セットTのトレーニング画像の特徴を抽出する画像特徴抽出ユニットと、(b29)QFTとzTとを選択し、画像特徴抽出ユニット(a24)が抽出した特徴を利用して最適品質係数の平均QFAおよび最適スケーリング係数の平均zAを決定する第1の品質認識パラメータ選択ユニットと、(c29)第1の品質認識パラメータ選択ユニット(b29)が選択したQFTおよびzTを利用して、トレーニング画像を、トランスクードされたトレーニング画像にトランスクードするトランスクードユニットと、(d29)トランスクードユニット(c29)が生成したトランスクードされたトレーニング画像の、トレーニング画像とトランスクードされたトレーニング画像との間の類似度を示す品質メトリックを決定する品質評価ユニットとを有する。

20

【0025】

第1の品質認識パラメータ選択ユニット(b29)はさらに、画質メトリックの平均QAを決定する計算手段を有する。トランスクードユニット(c29)は、(a31)スケーリング係数zTをトレーニング画像に適用することで、伸張およびスケーリングされたトレーニング画像を生成する伸張およびスケーリングユニットと、(b31)トランスクード品質係数QFTを伸張およびスケーリングされたトレーニング画像に適用することで、トランスクードされたトレーニング画像を生成する圧縮ユニットとを有する。品質評価ユニット(d29)は、(a32)リスケーリング係数zRを適用することで、トランスクードされたトレーニング画像を伸張およびスケーリングする第2の伸張およびスケーリングユニットと、(b32)視聴スケーリング係数zVを適用することで、トレーニング画像を伸張およびスケーリングする第3の伸張およびスケーリングユニットと、(c32)トランスクードされたトレーニング画像の品質メトリックを計算する品質メトリック計算ユニットとを有する。記品質メトリック計算ユニット(c32)は、トレーニング画像とトランスクードされたトレーニング画像との間の類似性を示す構造的なSIMilarity(SSIM)インデックスを決定する計算手段を有する。

30

【0026】

オンライントランスクード(d24)は、(a34)入力画像の特徴を抽出する画像特徴抽出ユニットと、(b34)第1のマトリックスに格納されている最適品質係数の平均QFAと、第2のマトリックスに格納されている最適スケーリング係数の平均zAとを、画像特徴抽出ユニット(a34)が抽出した入力画像の特徴と端末の特徴とを利用して取得する段階と、(c34)最適パラメータ予測ユニット(b34)が取得したQFAおよびzAを利用してトランスクード品質係数QFTおよびトランスクードスケーリング係数zTを決定して、第3のマトリックスに格納されている画質の平均QAを取得する第2の品質認識パラメータ選択ユニットと、(d34)第2の品質認識パラメータ選択ユニット

40

50

(c 3 4) が選択する Q F T および z T を利用して入力画像を出力画像にトランスコードするトランスコードユニットとを有する。

【0027】

コンピュータにより実行されると上に記載された方法の段階それぞれを実行するコンピュータ可読プログラムコード命令を格納するコンピュータ可読格納媒体が提供される。

【0028】

複数のコンピューティングノードが通信ネットワークにより接続されたシステムであつて、各ノードは、端末での表示用に入力画像を出力画像にトランスコードするべく、プロセッサとコンピュータ可読格納媒体とを含み、システムは、(a 3 6) 第1のノードのコンピュータ可読格納媒体に格納されているトレーニング画像を含むトレーニング画像セット T を含むトレーニング画像レポジトリと、(b 3 6) 入力画像のトランスコードの前に、各トレーニング画像について最高品質のトランスコードされたトレーニング画像を生成するべくトレーニング画像セット T 内のトレーニング画像のセットの最適品質のトランスコードを行することで生成されるトランスコードパラメータを含み、且つ、端末の特徴を満たすデータセットを生成する、第2のノードのコンピュータ可読媒体に格納されたコンピュータ可読命令を含むオフラインデータセット生成器と、(c 3 6) オフラインデータセット生成器 (b 3 6) が生成したデータセットを第3のノードのコンピュータ可読格納媒体に格納するデータセットレポジトリと、(d 3 6) 端末の特徴と入力画像の特徴を利用してデータセットレポジトリ (c 3 6) に格納されているデータセットからトランスコードパラメータを選択して、選択されたトランスコードパラメータを利用して入力画像を出力画像にトランスコードする、第4のノードのコンピュータ可読媒体に格納されているコンピュータ可読命令を含むオンライントランスコーダとを備えるシステムが提供される。10

【0029】

方法の段階 (b 1) は、入力画像の色数、画像ファイルサイズ S (I)、画像の幅 W (I)、および画像の高さ H (I) を含む、入力画像の特徴を抽出する段階を有する。

【0030】

このように、最適品質に近い品質での画像の複雑度の低いトランスコードシステムおよび方法が提供された。20

【0031】

本発明の更なる特徴および利点は、以下の例示目的に示される実施形態の記載を、添付図面を参照しながら読むことで明らかになる。30

【図面の簡単な説明】

【0032】

【図1】先行技術のマルチメディアメッセージシステム 100 の一例を示す。

【図2】本発明の一実施形態における端末上の表示用の出力画像へと入力画像をトランスコードするシステム 200 を示す。

【図3】図2のオフラインデータセット生成器 204 の機能コンポーネントを示す。

【図4】図2のオンライントランスコーダ 208 の機能コンポーネントを表示する。

【図5】本発明の実施形態における端末の表示用の出力画像へと入力画像をトランスコードする方法の各ステップを示すフローチャートを示す。40

【図6】図5の「データセットを生成する」ステップを示すフローチャートである。

【図7】図6の「平均値 Q F A および z A を計算する」ステップを示すフローチャートである。

【図8】図7の「Best Q F T および最良の z T を計算する」ステップを示すフローチャートである。

【図9】図6の「平均値 Q A を計算する」ステップを示すフローチャートである。

【図10】図9の「選択されたトレーニング画像に対応する品質メトリックを計算する」ステップを示すフローチャートである。

【図11】図5の「データセットからトランスコードパラメータを選択して、入力画像を

50

トランスコードする」ステップを示すフローチャートである。

【図12】本発明の実施形態の方法を示すのに利用される80のQF(I)に対応する第1のマトリックスの「スライス」を示す。

【図13】本発明の実施形態の方法を示すのに利用される80のQF(I)に対応する第2のマトリックスの「スライス」を示す。

【図14】本発明の実施形態の方法を示すのに利用される80のQF(I)に対応する第3のマトリックスの「スライス」を示す。

【発明を実施するための形態】

【0033】

本発明は、品質係数とスケーリング係数とをJPEGトランスコードにおいて組み合わせて、端末の解像度および画像のファイルサイズの制約を満たし、同時に、品質メトリックを最大化することに焦点をあてている。便宜上、本願で利用する用語集を以下に示す。

<用語集>

【0034】

H(D)は、端末の高さのことである。H(I)は、入力画像の、画像高さのことである。QAとは、トレーニング画像のサブセットの画質メトリックの平均であり、トレーニング画像と、トレーニング画像のサブセットにおけるトレーニング画像の最適品質のトランスコード中に生成された、それぞれ最高品質のトランスコードされたトレーニング画像との間の類似度で示される。QFAは、トレーニング画像のサブセットの最適品質係数の平均であり、トレーニング画像のサブセットにおけるトレーニング画像の最適品質のトランスコード中に利用されるトランスコード品質係数(QFT)の平均として決定される。QF(I)は、入力画像の符号化の忠実度を特徴付ける入力画像の品質係数である。QFTは、出力画像の符号化の忠実度を特徴付けるトランスコード品質係数である。S(D)は、端末がサポートする最大の画像ファイルサイズS(D)である。S(I)は、入力画像の画像ファイルサイズである。maxは、入力画像に対する出力画像の相対サイズを特徴付ける最大の相対画像サイズである。SSIMは、2つの画像の類似性を比較する際に利用される構造類似インデックス関数(Structural Similarity index function)のことである。Tは、画像トレーニングセットのことである。W(D)は、端末の幅である。W(I)は、入力画像の画像幅のことである。ZAは、トレーニング画像のサブセットの最適スケーリング係数の平均のことであり、トレーニング画像のサブセットにおけるトレーニング画像の最適品質トランスコード中に利用されるトランスコードスケーリング係数(zT)の平均として決定される。zTは、入力画像を出力画像にトランスコードすることに起因する画像の解像度の変化を特徴付けるトランスコードスケーリング係数である。zVは、出力画像の視聴条件による画像の解像度の変化を特徴付ける視聴スケーリング係数のことである。zRは、zVおよびzTの割合として決定される、リスケーリング係数のことである。

【0035】

本願は、JPEG画像のトランスコードに関する問題というコンテクストで、これら表記を利用する。Iは、JPEG圧縮画像のこととして、QF(I)、S(I)、W(I)、およびH(I)をそれぞれ、品質係数、画像ファイルサイズ(圧縮後)、画像幅、および画像高さとする。ここで、QFは、T. Lane, P. Gladstone, L. Ortiz, J. Boucher, L. Crocker, J. Minguillon, G. Phillips, D. Rossi、およびG. Weijersが「The independent JPEG group software release 6b」1998年で記載している独立JPEGグループ(the Independent JPEG Group)における定義にコンパイルされるものと仮定する。独立JPEGグループ(IJG)ソフトウェアおよび文書については、<http://www.ijg.org/>および<http://www.ijg.org/files/>を参照されたい。IJG-JPEGコーデックは、殆どの閲覧者および画像視聴者が利用しており、自由に利用できる事実上の標準的実装である。

【0036】

トランスコードされた画像は、一式の特徴を有する端末またはデバイス上に表示される

10

20

30

40

50

。端末およびデバイスという用語は、本願では交換可能な用語として用いられる。端末またはデバイスDでは、特徴には、端末がサポートする最大画像ファイルサイズS(D)、端末幅W(D)、および端末高さH(D)が含まれる。0 < z < 1を、アスペクトが保持されるスケーリング係数、あるいはズーム係数とする。JPEGトランスクード演算は、

(I, QFT, zT)と記述され、これは、トランスクードされた画像の符号化の忠実度および圧縮度(ひいてはファイルサイズ)を特徴付ける品質係数QFTと、トランスクードによる画像解像度の変化を特徴付けるトランスクードスケーリング係数zTとを、両方ともJPEG画像Iに適用することによって生じる圧縮された画像を返すことのできる関数である。品質係数(独立JPEGグループのライブラリにおける0-100スケールにおける)は、JPEG符号器で利用される量子化ステップを制御して、各周波数成分の解像度の低減度合いに影響を及ぼす。周波数成分に対して良好な解像度(小さい量子化ステップ)を利用することにより、高い品質係数は、画像を高い忠実度で符号化する。他方で、低い品質係数は、周波数成分に対して良好ではない解像度(大きな量子化ステップ)を利用して、しばしば高周波数成分(画像の詳細)を全て除去することにより、画像を低い忠実度で符号化する。品質係数が大きいということは、圧縮後のファイルサイズが大きいことを示す。従って品質係数は、圧縮度を制御して、究極的には再構成される画像が持つことを予期される忠実度を特徴付ける。

【0037】

JPEGトランスクード演算(I, QFT, zT)は、画像I、トランスクードパラメータQFT、およびzTが、端末の特徴により課される以下の制約の全てを満たす場合に、端末Dの上で実現可能として定義される。

$$\begin{aligned} S((I, QFT, zT)) &= S(D) \\ W((I, QFT, zT)) &= W(D) \\ H((I, QFT, zT)) &= H(D) \end{aligned}$$

【0038】

ここで、S(X)、W(X)、H(X)は、関数 が生成するトランスクードされた画像についての、画像のサイズ、画像の高さ、および画像の幅をそれぞれ表している。

【0039】

図2は、端末での表示用に入力画像を出力画像にトランスクードするシステム200を示す。システム200は、トレーニング画像レポジトリ202、オフラインデータセット生成器204、データセットレポジトリ206およびオンライントランスクード208を含む。マルチメディア画像のトランスクードを行う前に、オフラインデータセット生成器204は、トレーニング画像レポジトリ202に格納されているトレーニング画像セットTを利用して、データセットレポジトリ206に格納されているデータセットを生成する。Tは、トランスクードパラメータの生成に利用される画像およびデータセットに含まれる予測画質の複数の例を含む。オフラインデータセット生成器204およびトレーニング画像レポジトリ202は、データセットの決定のみに利用される。ひとたびデータセットが生成されると、オンライントランスクード208のみがMMSメッセージ処理に利用可能となる。オンライントランスクード208は、MMSメッセージの画像のトランスクードに利用されるトランスクードパラメータを選択するべく、データセットレポジトリ206に格納されているデータセット、端末の特徴および入力画像の特徴を利用する。

【0040】

本発明の実施形態のシステムは、CPUおよびコンピュータ可読媒体(メモリその他)を有する汎用または専用コンピュータを含むことができ、システムは、ファームウェア、またはファームウェアと専用コンピュータ(例えば専用コンピュータ用のファームウェア)の組み合わせで実装可能である。トレーニング画像レポジトリ202およびデータセットレポジトリ206各々は、コンピュータ可読格納媒体に格納される。オフラインデータセット生成器204およびオンライントランスクード208は、コンピュータ可読格納媒体に格納されたコンピュータ可読命令を含む。

【0041】

10

20

30

40

50

図3は、オフラインデータセット生成器204の機能コンポーネントを示し、図4は、オンライントランスクーダ208の機能コンポーネントを示す。これら機能コンポーネントは、ファームウェアその他、コンピュータ可読格納媒体に格納されたコンピュータソフトウェアコードを含むユニットを含む。これらコンピュータ可読媒体は、本発明の実施形態の方法のステップを実行して、システム200の機能コンポーネントを形成する命令を格納しており、コンピュータメモリ、DVD、CD-ROM、フロッピー（登録商標）等を含んでよい。またはシステムが、通信ネットワークが接続する複数のコンピューティングノードを含み、各ノードが、後述するように、端末での表示用に入力画像を出力画像にトランスクードするために、プロセッサとコンピュータ可読格納媒体を含んでよい。

【0042】

10

図3が示すオフラインデータセット生成器204は、コンピュータ可読格納媒体に格納されておりプロセッサにより実行されるコンピュータ可読命令を含む。オフラインデータセット生成器204は、画像特徴抽出ユニット302、第1の品質認識パラメータ選択ユニット304、トランスクードユニット306、および品質評価ユニット308を含む。画像特徴抽出ユニット302は、トレーニング画像Lを取得して、第1の品質認識パラメータ選択ユニット304に提供する、トレーニング画像を符号化する際の忠実度を特徴付ける品質係数QF(L)、画像のファイルサイズS(L)、画像の幅W(L)、および画像の高さH(L)を含むトレーニング画像Lの特徴を抽出する。S(L)は、トレーニング画像の符号化中に実行される圧縮後に取得されるファイルサイズであり、画像ファイルサイズと称される。トレーニング画像Lの特徴に加えて、第1の品質認識パラメータ選択ユニット304は、品質評価ユニット308が計算するパラメータQF(I)、*smax*およびzVおよび品質メトリック、および、トランスクードユニット306が生成するトランスクードされたとトレーニング画像を利用する。zVは、出力画像の視聴条件による画像の解像度の変化を特徴付ける視聴スケーリング係数であり、*smax*は、トランスクード後に得られる画像の相対サイズを特徴付ける。

【0043】

20

*smax*の計算を含むこれらパラメータについては、本明細書で後述する。第1の品質認識パラメータ選択ユニット304は、任意の端末特徴に対応する最適トランスクードパラメータと、トレーニング画像Lの特徴とのセットを決定する。最適トランスクードパラメータは、トレーニング画像Lについての最高品質のトランスクードされた画像を生成する最適品質トランスクードに対応する品質係数およびスケーリング係数を含む。第1の品質認識パラメータ選択ユニット304は、この最高品質のトランスクードされた画像の品質メトリックも決定する。これら最適トランスクードパラメータおよび品質メトリックをトレーニング画像セットTの任意のトレーニング画像のサブセットに対して利用することで、該サブセットの最適品質係数の平均QFA、最適スケーリング係数の平均zA、および画質メトリックの平均QAを計算することができる。このトレーニング画像のサブセットは、例えば任意の品質係数を有するトレーニング画像に対応していてよい。第1の品質認識パラメータ選択ユニット304は、画質メトリックの平均QAおよび最適トランスクードパラメータの平均を決定するコンピュータ手段を提供するプロセッサ上で実行されるコンピュータ可読命令を含むプロシージャまたは機能を含む。

30

【0044】

40

第1の品質認識パラメータ選択ユニット304は、各反復でトランスクード品質係数QFTおよびトランスクードスケーリング係数zTを選択する反復法（後述する）を利用する。第1の品質認識パラメータ選択ユニット304のzTおよびQFTは、入力トレーニング画像Lを処理するトランスクードユニット306に利用される。そしてトランスクードユニット306は、第1の伸張およびスケーリングユニット310および圧縮ユニット312を利用する。第1の伸張およびスケーリングユニット310は、トレーニング画像Lを入力として受信して、第1の品質認識パラメータ選択ユニット304が生成したzTを適用して、トレーニング画像Lの伸張およびスケーリングを実行する。第1の伸張およびスケーリングユニット310の出力は、第1の品質認識パラメータ選択ユニット304

50

が生成するトランスクード品質係数 Q_{FT} を適用して、品質評価ユニット 308 の入力に対して提示されるトランスクードされたトレーニング画像 J を生成する圧縮ユニット 312 で利用される。品質評価ユニット 308 は、第 2 の伸張およびスケーリングユニット 314 および第 3 の伸張およびスケーリングユニット 318、および、品質メトリック計算ユニット 316 を含む。トランスクードユニット 306 の出力であるトランスクードされたトレーニング画像 J 、および、トレーニング画像 L は、両方とも、品質評価の前にスケーリングされる。トレーニング画像 L とそのトランスクードされたバージョンとを比較する際の品質メトリックを定義して、出力画像の視聴条件による画像解像度の変化を特徴付ける視聴スケーリング係数 z_V 用のトランスクードされたトレーニング画像 J (トランスクードスケーリング係数 z_T を利用してトランスクードされて生成される) は、 $Q_{zV}(L, J) = SSIM(R(L, zV), R(J, zV / zT))$ として表すことができる。
ここで $SSIM$ は、2つの画像の類似度を比較する際に利用される構造類似度インデックスであり、Zhou Wang、Alan Conrad Bovik、Hamid Rahim Sheikh および Eero P. Simoncelli による「Image quality assessment: from error visibility to structural similarity」なる研究論文、IEEE Transactions on Image Processing、第13巻、4号、600-612ページ、2004年4月で定義されている。 $R(I, z)$ は、 I を伸張して、スケーリング係数 z を利用してスケーリングする演算子である。この定義および図 3 によると、画像解像度が等しい場合には、 $z_V = z_T * z_R$ という等式が満たされる。ここでは、品質を比較するときには原画像の解像度を上げたくないでの $z_V < 1$ であることから、端末の特徴が課す制約を満たすために $z_T < 1$ であることが必須である。視聴スケーリング係数 z_V ($z_T < z_V < 1$) が制御する出力画像の視聴条件は、トランスクード演算結果をユーザが知覚する際に重要な役割を果たす。例えば、 $z_V = 1$ である場合、原入力トレーニング画像 L の解像度で画像間を比較すると、 $z_R = 1 / z_T$ となる。

【0045】

$z_T < z_V < 1$ である場合、原トレーニング画像の解像度とトランスクードされたトレーニング画像 J の解像度との間の解像度で画像間を比較する。品質メトリック計算ユニット 316 は、プロセッサ上で実行されるコンピュータ可読命令を含み、 $SSIM$ インデックスを決定する計算手段を提供するプロシージャまたは関数を含む。第 2 の伸張およびスケーリングユニット 314 は、スケーリング係数 z_R を利用して画像 J' を生成することによりトランスクードされたトレーニング画像 J をスケーリングして、第 3 の伸張およびスケーリングユニット 318 は、視聴スケーリング係数 z_V を利用して画像 L' を生成することにより、入力トレーニング画像 L をスケーリングする。品質メトリック計算ユニット 316 は、画像 J' 、 L' を利用して、これら 2 つの画像を利用して品質メトリック $Q_{zV}(L', J')$ を生成することで、 $SSIM$ インデックスを決定して品質認識パラメータ選択ユニット 304 に提示する。

【0046】

これら機能コンポーネント日々、および、オンラインデータセット生成器 204 が利用するトレーニング画像レポジトリ 202、および、オンラインデータセット生成器 204 およびオンライントランスクード 208 が両方とも利用するデータセットレポジトリ 206 は、コンピュータ格納媒体（例えばメモリ、ディスク、DVD、CD-ROM、フロッピー（登録商標）等）を含む。または、トレーニング画像レポジトリ 202 およびデータセットレポジトリ 206 は、インターフェースを介して他のシステムコンポーネント（例えば図 2 に示すもの）に接続されたコンピュータ可読媒体およびコンピュータ格納媒体に格納されているプロセッサ実行命令を備えた独立したコンピュータを含んでもよい。ここでの接続は、ローカルなものであっても、遠隔位置にあるトレーニング画像レポジトリ 202 またはデータセットレポジトリ 206 用にインターネットを挟んでコンピュータと通信可能なものであってもよい。

【0047】

図 4 で表示されているオンライントランスクード 208 は、コンピュータ格納媒体に格納されており、プロセッサが実行するコンピュータ可読命令を含む。オンライントランス

コーダ 208 は、画像特徴抽出ユニット 302、最適パラメータ予測ユニット 402、第 2 の品質認識パラメータ選択ユニット 404、およびトランスクードユニット 306 を含む。画像特徴抽出ユニット 302 およびトランスクードユニット 306 は、オフラインデータセット生成器 204 で利用されるものと全くおなじである。ここでの唯一の差異は、オフラインデータ生成器 204 のコンテクストで利用されたトレーニング画像のトランスクードと違って、MMS メッセージからの入力画像をトランスクードする、ということである。画像特徴抽出ユニット 302 は、トランスクードする入力画像 I の特徴を抽出して、抽出した特徴を最適パラメータ予測ユニット 402 および第 2 の品質認識パラメータ選択ユニット 404 に渡す。

【0048】

10

これら特徴は、入力画像の符号化における忠実度を特徴付ける品質係数 QF(I)、画像ファイルサイズ S(I)、画像の幅 W(I)、および画像の高さ H(I) を含む。入力画像のこれら特徴を利用して、最適パラメータ予測ユニット 402 は、データセットレポジトリ 206 に格納されているデータセットから最適トランスクードパラメータの平均を取得して、これらを第 2 の品質認識パラメータ選択ユニット 404 に渡す。画像のトランスクードは、トランスクードされた画像が端末の特徴が課す制約を満たしたときに停止される反復プロセスである。最適パラメータ予測ユニット 402、第 2 の品質認識パラメータ選択ユニット 404、および、トランスクードユニット 306 は、この反復プロセスで互いに協同する。第 2 の品質認識パラメータ選択ユニット 404 は、トランスクードユニット 306 が利用するトランスクードパラメータを生成する。第 2 の品質認識パラメータ選択ユニット 404 は、最適パラメータ予測ユニット 402 が取得する情報と、トランスクードユニット 306 が生成するトランスクードされた画像の特徴とを利用する。最適パラメータ予測ユニット 402 から取得されるパラメータの値は、トランスクード演算を行うトランスクードユニット 306 に渡される。トランスクードされた画像は、出力画像のサイズをチェックして、反復をとめるか否かを判断する第 2 の品質認識パラメータ選択ユニット 404 に転送される。この反復の「ステータス」に関する決定は、最適パラメータ予測ユニット 402 へと渡される。反復を続ける場合には、最適パラメータ予測ユニット 402 は、データセットからトランスクードパラメータの新しい値を取得して、これらを第 2 の品質認識パラメータ選択ユニット 404 に転送する。反復を続けない場合には、出力画像の予測品質をデータセットから取得して、この値を、出力画像（最後の反復で生成される）および最後の反復で利用される QFT および ZT の値とともに出力する第 2 の品質認識パラメータ選択ユニット 404 に転送する。出力画像は、データセットに記録されている最適値の平均に近い値でパラメータをトランスクードすることで生成されたものなので、「最適に近い」と称される。本明細書で後述するシステムの実験的分析によって、この「最適に近い」画像の品質メトリックが、実際に、最適品質トランスクードから得られる出力画像の品質メトリックに近いことが証明された。

【0049】

30

端末での表示用に入力画像を出力画像にトランスクードする方法の各ステップを、図 5 に示すフローチャート 500 を参照して説明する。方法は 2 つのステップを含む。ボックス 504 に記載する方法のステップは、入力画像のトランスクードの前に一度だけ行われる。ボックス 506 に記載する方法のステップは、入力画像のトランスクード中に実行される。マルチメディアストリームにおける画像のトランスクードに利用される場合には、プロシージャは、マルチメディアストリームの一部である複数の入力画像に対してボックス 506 の処理を行うことができる。

【0050】

40

開始されると(502)、プロシージャ 500 は、トレーニング画像レポジトリ 202 に格納されているトレーニング画像セットを利用して、端末特徴およびトレーニング画像の特徴の様々な組み合わせに対応する最適トランスクードパラメータの平均および画質の平均を含むデータセットを生成する(ボックス 504)。破線のボックスは、この処理が、MMS 画像のトランスクード処理の前に、オフラインデータセット生成器 204 により

50

オフラインで行われることを示している。データセットの生成の後、MMSメッセージ（オンライントランスクーダ208で行われ、プロシージャ500が、データセットからトランスクードパラメータを選択する（ボックス506））の処理中に行われ、入力画像をトランスクードする。プロシージャ500は、ボックス506により実行される処理の最後に終了し、退出する（ボックス508）。

【0051】

図5の「データセットを生成する」ステップ（ボックス504）を、図6に示すフローチャート600を参照しながら詳述する。プロシージャ600の目的は、固定間隔で変化する値を有するQF(I)、smax、およびzVの様々な組み合わせにとって最適なトランスクードパラメータおよび画質を生成することである。上述したように、smaxパラメータは、トランスクードされた画像（トレーニング画像または入力画像）のサイズに比した場合の、トランスクード後に得られた画像（例えばトランスクードされたトレーニング画像および出力画像）の相対サイズを特徴付ける。特定の端末が与えられたときのトランスクード後の画像の最大許容相対サイズとして定義され、 $smax = min((S(D)/S(I)), 1)$ と表される。ここでIはトランスクードする画像であり、Dは、トランスクードされた画像を表示する端末である。
10

【0052】

開始すると（ボックス602）、プロシージャ600は、3つのパラメータQF(I)、smax、およびzVの最初の組み合わせを選択する（ボックス604）。プロシージャ600はその後、この組み合わせに最適なトランスクードパラメータの平均値QFAおよびzAを計算する（ボックス606）。QFAおよびzAの値は、第1および第2のマトリックスにそれぞれ格納される（ボックス608）。この反復で利用されるQF(I)、smax、およびzVによりインデックスされる第1および第2のマトリックスの位置を利用してこれら値を格納する。次に画質メトリックの平均QAを計算して（ボックス609）、第3のマトリックスの、この反復で利用されるQF(I)、smax、およびzVの値によりインデックスされる位置に格納する（610）。そしてプロシージャ600は、検討が必要なQF(I)、smax、およびzVの組み合わせがまだ他にあるかをチェックする（ボックス611）。チェック結果が肯定的である場合には、プロシージャ600はボックス611から「YES」へ行き、次のQF(I)、smax、およびzVの組み合わせを選択して（ボックス612）、ボックス606の始めにループして戻る。チェック結果が否定的である場合には、プロシージャ600はボックス611から「NO」へ行く（ボックス614）。

【0053】

図6の方法のステップ「平均QFAおよびzAを計算する（ボックス606）」を、図7に示すフローチャート700を参照して説明する。

【0054】

開始すると（ボックス702）、プロシージャ700は、プロシージャ600で選択したパラメータの組み合わせに対応し、端末の特徴が課す制約を表すパラメータzVおよびsmaxを取得する（ボックス702）。そしてプロシージャ700は、品質係数QF(I)を有するトレーニング画像セットTの各トレーニング画像の処理の反復を開始する準備をする。Tからの第1のトレーニング画像がLに格納される（ボックス704）。Lに格納される画像は、画像Lとも称される。そしてプロシージャ700は、Lに格納されているトレーニング画像の特徴を抽出して（ボックス706）、パラメータを初期化して（ボックス708）、Lに格納されているトレーニング画像のトランスクードに利用されるQFTおよびzTの、端末の制約を満たす最良の値を計算する（ボックス710）。QFTの最良の値であるBestQFT、および、zTの最良の値であるBestzTは、最高品質のトランスクードされたトレーニング画像を生成する、トレーニング画像セットTのトレーニング画像の最適品質トランスクードに対応するトランスクードパラメータであり、且つ、zT zV、および、S((L, QFT, zT)) smaxという制約において取得される端末の特徴を満たす。これらQFTおよびzTの最良の値を格納して
40
50

(ボックス 714) 、プロシージャ 700 が、品質係数が Q F (I) に等しいトレーニング画像が未だ他に T にあるかをチェックする (ボックス 716) 。チェック結果が肯定的である場合には、プロシージャ 700 はボックス 716 から「 Y E S 」へ行き、 T から次の画像を L に格納して、ボックス 706 のエントリにループして戻る (ボックス 712) 。チェック結果が否定的である場合には、プロシージャ 700 はボックス 716 から「 N O 」へ行き、品質係数が Q F (I) に等しいトレーニング画像のサブセットから平均値 Q F A および z A を計算して (ボックス 718) 、これら値を戻し (ボックス 720) 、退出する (ボックス 722) 。

【 0055 】

最適トランスコードパラメータの平均値は、トレーニング画像セット T のトレーニング画像を利用して、 $Q F A = (1/N) \sum_{T \in Q F I} Q F T_{best}(L)$ 、および、 $z A = (1/N) \sum_{T \in Q F I} z T_{best}(L)$ として計算される。ここで $T Q F I$ は、品質係数が Q F (I) に等しいトレーニング画像を含む T のサブセットであり、 N は、このサブセットの濃度であり、 $Q F T_{best}(L)$ および $z T_{best}(L)$ は、前述したプロシージャ 600 のボックス 614 の L 番目のトレーニング画像について格納されている Best Q F T および Best z T の値である。

【 0056 】

また、 Q F A および z A の計算には他の反復プロシージャを利用することもできる。各反復中に、各トレーニング画像について別個に Best Q F T および Best z T を格納する代わりに、 $Best Q F T_{acc} = Best Q F T_{acc} + Best Q F T$ 、および、 $Best z T_{acc} = Best z T_{acc} + Best z T$ とすることで、 Best Q F T および Best z T の値を変数 Best Q F T acc および Best Z t acc に累積することができる。 Best Q F A および Best z T の N 個の値が累積されたので、 $Q F A = Best Q F T_{acc} / N$ 、および、 $Z A = Best z T_{acc} / N$ となる。

【 0057 】

図 7 の方法のステップ「 Best Q F T および Best z T を決定する」 (ボックス 710) を、図 8 のフローチャート 800 を参照して詳述する。プロシージャ 800 は、段階的に z T および Q F T を変化させて、これらパラメータの全ての可能な組み合わせを生成して、どのパラメータの組み合わせが、与えられているトレーニング画像 (L に格納されている) についてトランスコードされたトレーニング画像 (J に格納されている) にとって最適品質を生じるかを判断する。開始すると (ボックス 802) 、プロシージャ 800 は、トランスコードに利用可能な他の z T 、 Q F T の対が存在するかどうかをチェックして反復を開始する (ボックス 804) 。このような対が存在するためには、 z T が z V 以下である必要がある点に留意されたい。さらなる z T 、 Q F T の対が可能である場合、プロシージャ 800 はボックス 804 から「 Y E S 」へ行き、次の z T , Q F T の対を取得する (ボックス 806) 。与えられているトレーニング画像を、 J に格納されているトランスコードされたトレーニング画像へとトランスコードする (ボックス 808) 。トランスコードされたトレーニング画像は、以下の説明では J と称される。次に、トランスコードされているトレーニング画像 J のサイズをチェックする (ボックス 810) 。 s m a x が課すサイズの制約を満たさない場合には、プロシージャ 800 はボックス 810 から「 N O 」へ行き、ボックス 804 のエントリへとループして戻る。満たす場合には、プロシージャはボックス 810 から「 Y E S 」へ行き、 z V をスケーリング係数として利用して、 L のトレーニング画像を伸張して、 L' に格納される画像を生成して、 L' と称する (ボックス 812) 。次のステップで、プロシージャ 800 は、 z R をスケーリング係数として利用して、トランスコードされたトレーニング画像 J を伸張して、 J' に格納される画像を生成して、 J' と称する (ボックス 814) 。そしてプロシージャ 800 は、 S S I M インデックスを利用して品質メトリック Q z V (L' 、 J') を計算する (ボックス 816) 。 Q z V (L' 、 J') の値が高いほど、トランスコードされたトレーニング画像の品質も良好である。次に、 Q z V (L' 、 J') を、今まで計算したものの中で最高の品質メトリックを含む Best Q に格納されている値と比較する (ボックス 818) 。 Q z V

10

20

30

40

50

(*L'*、*J'*) が Best Q より高い場合には、プロシージャ 800 はボックス 818 から「YES」へ行き、変数 Best Q、Best QFT、および Best zT に *QzV* (*L'*、*J'*)、*QFT*、および *zT* の値をそれぞれ記録して(ボックス 820)、ボックス 804 のエントリループして戻る。*QzV* (*L'*、*J'*) が Best Q より高くない場合には、プロシージャ 800 はボックス 820 の処理をスキップして、単にボックス 804 のエントリにループして戻る。ボックス 804 で行ったテスト結果が偽である場合には、チェックすべき *zT*、*QFT* 対がもうないことになるので、プロシージャはボックス 804 から「NO」へ行き、Best Q、Best QFT、および Best zT に格納されている値を返し(ボックス 822)、退出する(ボックス 824)。

【0058】

10

図 6 のステップ「平均 QA を計算する」(ボックス 609)を、図 9 のフローチャート 9 を参照して詳述する。開始すると(ボックス 902)、プロシージャ 900 は、トレーニング画像セット T から品質係数が *QF* (*I*) に等しい第 1 の画像を選択する(ボックス 904)。次に、累積品質係数を 0 に初期化する(ボックス 906)。次にプロシージャ 900 は、SSIM インデックスを決定する関数を利用して、選択されたトレーニング画像に対応する品質メトリックを計算して(ボックス 908)、累積品質メトリックをこれに基づいて更新する(ボックス 910)。次に、T の品質メトリック *QF* (*I*) を有する画像がまだあるかをチェックする(ボックス 912)。まだある場合には、この品質係数を有する別の画像を T から選択して(ボックス 918)、プロシージャ 900 は、ボックス 908 のエントリループして戻る。もうない場合には、プロシージャ 900 は、ボックス 912 から「NO」へ行き、累積品質メトリックから *QA* の値を計算して(ボックス 914)、退出する(ボックス 916)。*QA* は、 $QA = \text{累積品質メトリック} / N$ として計算される。*N* は、品質係数が *QF* (*I*) に等しいトレーニング画像のサブセットのトレーニング画像数である。

【0059】

20

図 9 のステップ「選択されたトレーニング画像の品質メトリックを計算する」(ボックス 908)を、図 10 のフローチャート 1000 を参照して詳述する。開始すると(ボックス 1002)、プロシージャ 1000 は、それぞれ第 1 および第 2 のマトリックスの、*QF* (*I*)、*smax*、および *zV* の値でインデックスされた位置から、格納されている *QFO* および *zO* の値を取得する(ボックス 1004)。次にトランスコードパラメータを決定して、*QFT* を *QFO* に設定して、*zT* を *zO* に設定する(ボックス 1006)。次に画像を、これらの *QFT* および *zT* の値を利用してトランスコードする(ボックス 1008)。次に、こうしてトランスコードされたトレーニング画像のファイルサイズが端末にとって大きすぎないかをチェックする(ボックス 1010)。大きすぎる場合には、プロシージャ 1000 は、ボックス 1010 から「YES」へ行き、*smax* の値を、第 1 および第 2 のマトリックスで利用される次に小さい値に設定して(ボックス 1014)、ボックス 1004 のエントリにループして戻る。大きすぎない場合には、プロシージャはボックス 1010 から「NO」へ行き、SSIM インデックスを決定する関数を適用して、トランスコードされた画像の品質メトリックを計算して(ボックス 1012)、退出する(ボックス 1016)。上述したように、トレーニング画像とトランスコードされたトレーニング画像とを等しい解像度で比較するには、両方の画像を適切にスケーリングする必要がある。トレーニング画像は *zV* を利用してスケーリングし、トランスコードされたトレーニング画像は *zR* を利用してスケーリングし、ここで *zV* および *zR* は、 $zV = zT * zR$ を満たす。

30

【0060】

40

図 5 のステップ「データセットからトランスコードパラメータを選択して、入力画像をトランスコードする」(ボックス 506)を、図 11 のフローチャート 1100 を参照して詳述する。開始すると(ボックス 1101)、プロシージャ 1100 は、端末の特徴を取得して(ボックス 1102)、入力画像の特徴を取得する(ボックス 1103)。*smax* および *zV* の値を、 $smax = \min((S(D) / S(I)), 1)$ 、および、*z*

50

$V = \min((W(D) / W(I)), (H(D) / H(I)), 1)$ を利用して計算する（ボックス 1104）

【0061】

smax のこの値は、任意の端末 D の画像 I の最大許容可能な相対サイズに対応し、選択された zV は、端末がサポートする最大解像度を生じる。

【0062】

トランスコードパラメータを決定して、トランスコードされた画像にとって最適に近い画質を得る反復は、第 1 および第 2 のマトリックスの $QF(I)$ 、 smax 、および zV でインデックスされる位置の QFO および ZO の値を読み出すことから始まる（ボックス 1106）。次に QFT を QFO に設定して、 zT を ZO に設定する（ボックス 1108）。そして入力画像を、 QFT に等しい品質係数およびスケーリング係数 zT でトランスコードする（ボックス 1112）。次に、トランスコードされた画像のファイルのサイズをチェックする（ボックス 1114）。サイズが、端末の特徴が課す制約を満たすには大きすぎる場合には、プロシージャ 1100 はボックス 1114 から「YES」へ行き、第 1 および第 2 のマトリックスの smax の次に小さい値を特定して、ボックス 1106 のエントリにループして戻る。サイズが大きすぎない場合には、プロシージャ 1100 は、ボックス 1114 から「NO」へ行き、予測画質、トランスコードで利用された QFT および zT の値、および最後にトランスコードされた画像を戻し（ボックス 1116）、退出する（ボックス 1118）。予測画質は、 $QF(I)$ 、 smax 、および zV をインデックス変数として利用して第 3 のマトリックスから簡単に取得される。

10

20

【0063】

本発明の実施形態の方法の重要な特徴は、画像の予測品質メトリックを予め計算して格納することである。品質メトリックの決定に利用される SSIM インデックスを計算するとコストが高くなるので、それを、インデックスが量子化されたパラメータであるアレイに予め計算する必要がある。これら量子化されたパラメータは、量子化された zV 、量子化された $QF(I)$ 、および量子化された smax を含む。パラメータの量子化により、互いに近いパラメータの値を同じグループにして、これらを単一の値で表すことができるようになる。予め計算するために、画像トレーニングセット T の画像を利用する。T 内の各トレーニング画像 L について、固定間隔で変化する異なる QFT および zT を利用して多数の変換を適用する。結果生じる、トランスコードされたトレーニング画像用の画像のファイルサイズおよび品質メトリックを記録した。各トランスコードについて、特徴ベクトル t（トレーニング画像 L、 $QF(L)$ 、 $W(L)$ 、 $H(L)$ 、 $S(L)$ 、 QFT 、 zT 、視聴スケーリング係数 zV を含む）、結果生じるトランスコードされたトレーニング画像のファイルサイズ S ((L, QFT, zT))、および、トランスコードされたトレーニング画像の品質メトリック $QzV(L, T(I, QFT, zT))$ を、視聴スケーリング係数 zV で表される視聴条件で形成する。これら全てのベクトルに、トランスコードされた画像セット U を形成させる。

30

【0064】

量子化された $QF(I)$ 、量子化された QFT 、および量子化された zV の SSIM インデックスに基づく品質予測は、 $SSIM - Q = (1/M) \sum_t u \cdot SSIM(R(L(t), zV(t)))$ で表される。ここで U' は、パラメータが量子化セル、量子化された $QF(I)$ 、量子化された QFT 、量子化された zT 、および量子化された zV に収まるトランスコードされた画像セット U の全てのトランスコードされたトレーニング画像のうちのサブセットであり、M はセット U' の濃度であり、L(t) はトレーニング画像 L であり、 $QFT(t)$ はトランスコード QFT であり、 $zT(t)$ は、ベクトル t から適用されたトランスコードスケーリング係数である。

40

【0065】

次に、画像トランスコードにおいて本発明の実施形態の方法およびシステムを利用する一例を示す。Steven Pigeon、Stephane Coulombeによる「Computationally efficient al

50

gorithms for predicting the file size of JPEG images subject to changes of quality factor and scaling」という研究論文、Proceedings of the 24th Queen's Biennial Symposium on Communications, Queen's University, Kingston, Canada, 2008で記載されている多数の画像から、利用するデータセットが得られる。マルチメディアアプリケーションからサンプルされる通常のJPEG画像の大きなデータベースは利用できないので、人気のあるウェブサイトから画像を抽出するためにクローラを開発した。アセンブルされたトレーニング画像セットは、数万というJPEGファイルを含む。破損したファイルおよび全てのメタデータ(Exif)を除去した。第1、第2、および第3のマトリックスを生成するべく、<http://www.imagemagick.org/>の「ImageMagick command-line tools」に記載されているImageMagickのコマンドラインツールを利用して各トレーニング画像に複数のトランスクード演算を行った。80のQF(I)に対応する各マトリックスの「スライス」を図12-14に示す。このスライスは、zVおよびsmaxでインデックスされる2次元マトリックスである。これらマトリックスのスライスがここで提供する例で利用される。10

【0066】

$S(D) = 30500$ 、 $W(D) = 640$ 、 $H(D) = 480$ 、および入力画像I、lenaであり、 $S(I) = 43266$ 、 $W(I) = 512$ 、 $H(I) = 512$ 、および $QF(I) = 80$ であるデバイスを考える。画像lenaに関する情報に関しては<http://sipi.usc.edu/database/database.cgi?volume=misc&image=12>, <http://www.cs.cmu.edu/~chuck/lennapg/>を参照されたい。ここで、ステップ1では、 $smax = \min(30500 / 43266, 1) = 0.7$ であり、 $zV = \min(640 / 512, 480 / 512, 1) = 90\%$ を計算する。ステップ2では、図12に示す第1のマトリックススライスを利用して、 $QFT = 第1のマトリックススライス[0.7, 90\%] = 70$ (69.3をパラメータの解像度に丸めた値)とする。第1のマトリックススライスは、 $QF(I) = 80$ に対応する二次元マトリックスである点に留意されたい。この二次元マトリックスの行が $smax$ の値に対応しており、列が zV の値に対応している。図13に示す第2のマトリックススライスを利用して、 $zT = 第2のマトリックススライス[0.7, 90\%] = 80\%$ となる。第2のマトリックススライスは、 $QF(I) = 80$ に対応する二次元マトリックスである点に留意されたい。この二次元マトリックスの行が $smax$ の値に対応しており、列が zV の値に対応している。ステップ3では、 QFT に等しい品質係数、および zT に等しいスケーリング係数で画像をトランスクードする。トランスクードした後で、入力画像に対して相対画像ファイルサイズが0.53である出力画像が得られる。ステップ4では、トランスクードされた画像が、画像ファイルサイズに対して端末の特徴が課す制約が満たされるので($0.53 < 0.70$)、ステップ5に進む。ステップ5で、最適に近いパラメータ $QFO = 70$ および $zO = 80\%$ 、これらパラメータ($I, 70, 80\%$)を利用して得られた出力画像、図14が示す第3のマトリックススライス(第三のマトリックススライス[[0.7, 90%] = 0.84])から読み出されたトランスクードされた画像の予測画質 = 0.84を返す。20

【0067】

入力画像に最適品質トランスクードを行うことにより、0.85という画質が得られる。従ってlenaについては、最適に近い品質(0.84対0.85 1%の誤り)が得られた。シミュレーション実験を多数の他の画像について繰り返し、各実験結果が非常に小さい誤りとなった。30

【0068】

本発明の実施形態は、(1)入力画像を最適品質でトランスクードする方法から得られるものに近い高品質の出力画像が得られ、(2)最適品質のトランスクードで必要となる時間よりもとても少ないトランスクード時間でよく、(3)上記した2つの利点から、匹敵する品質の出力画像を生成しつつ、画像が維持することのできる最適品質のトランスクードシステムと比してより高いMMSメッセージレートが得られ、(4)システムが制約を満たす解決法を見つけられない頻度を示すインジケータとなる誤りレートを任意の所望40

の低い値に設定することができるようになる、といった利点を有する。

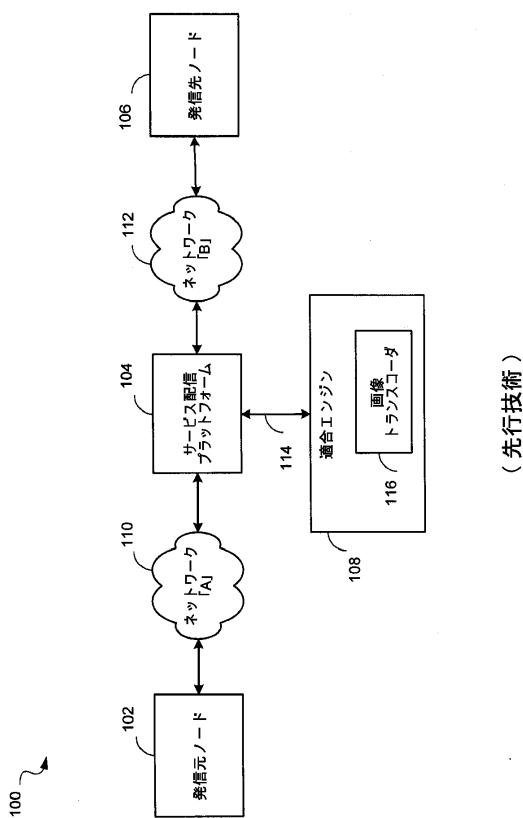
【0069】

本発明の具体的な実施形態を詳述してきたが、記載された実施形態は例示を意図しており限定は意図していない点に留意されたい。図面および明細書に示した実施形態には、添付請求項の範囲内で、且つ、本発明の広義の範囲を逸脱せずに、様々な変更例および変形例が可能である。例えば、本書類で記載したSSIM以外の品質メトリックを利用して画質を計測することもできる。本願で提示した説明は主にJPEG画像に関してのものであったが、本発明の実施形態の方法およびシステムは、他のフォーマットで符号化されたデジタル画像の処理に適用することもできる（例えば、GIF(Graphics Interchange Format)およびPNG(Portable Network Graphics)）。GIF画像の場合には、入力画像を出力画像にトランスクードする方法およびシステムにおいて、画像の色数をQF(I)、QFT、およびQFAの代わりに用いる必要が出てくるだろう。
10

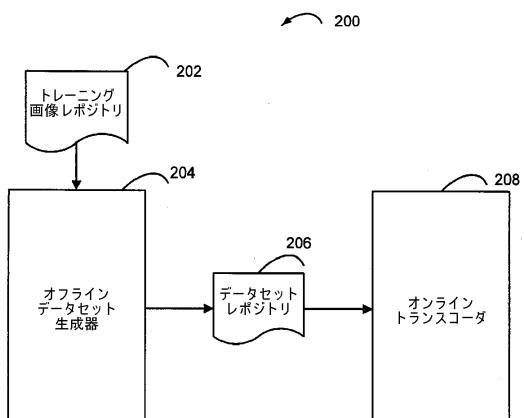
【0070】

本発明の実施形態を詳述したが、当業者であれば、これらの実施形態の変形例および変更例を、以下の請求項の範囲内で想到することができる。

【図1】

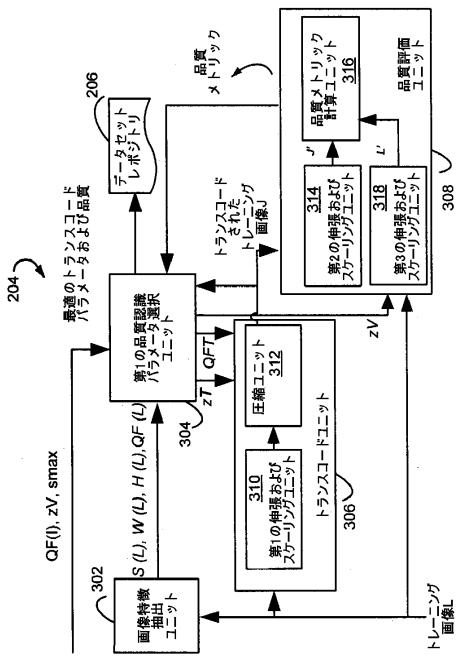


【図2】

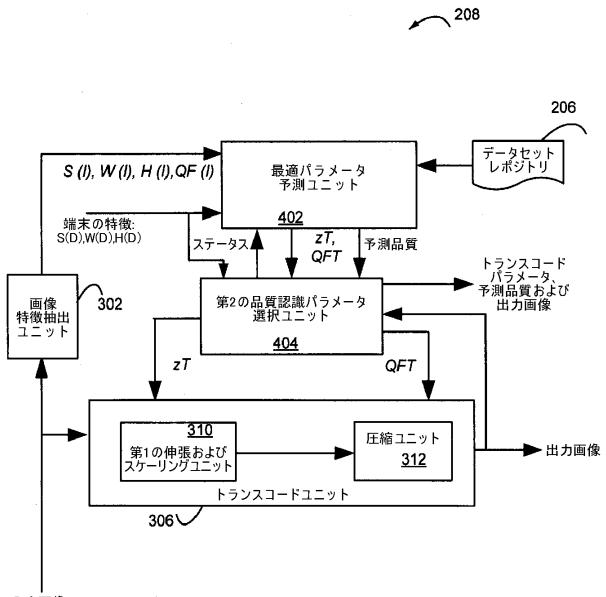


(先行技術)

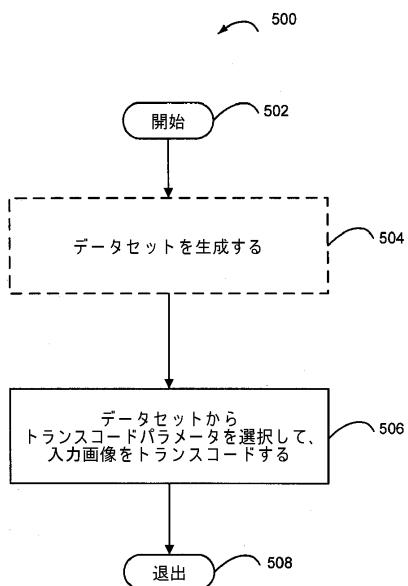
【図3】



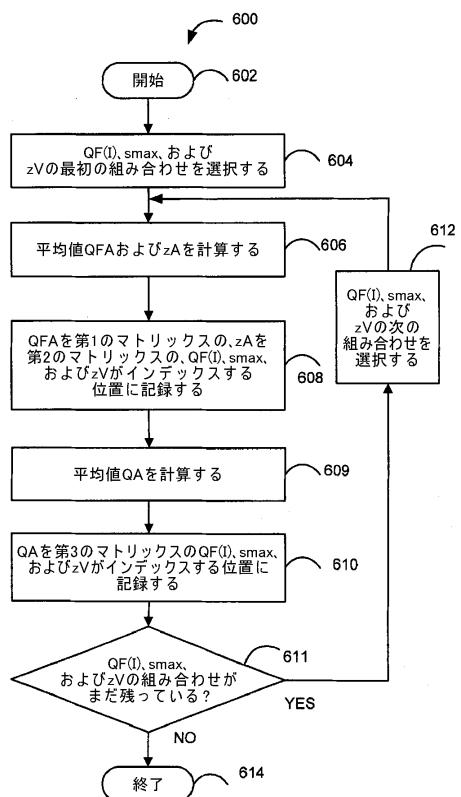
【図4】



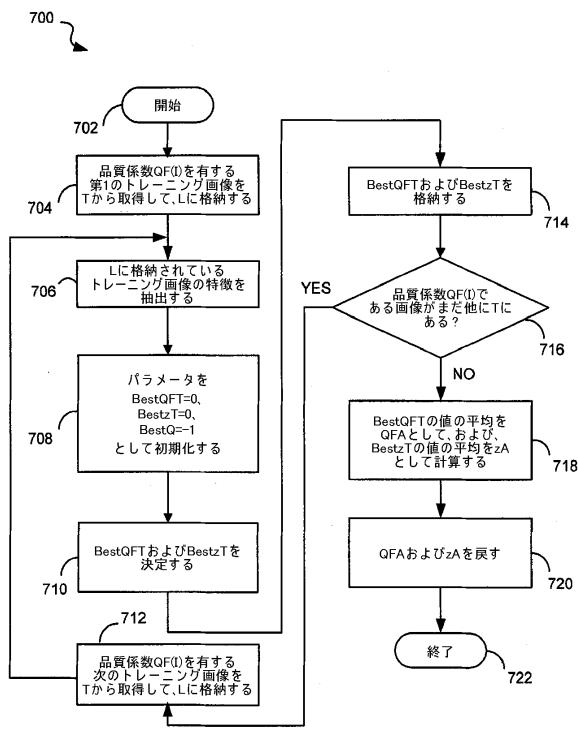
【図5】



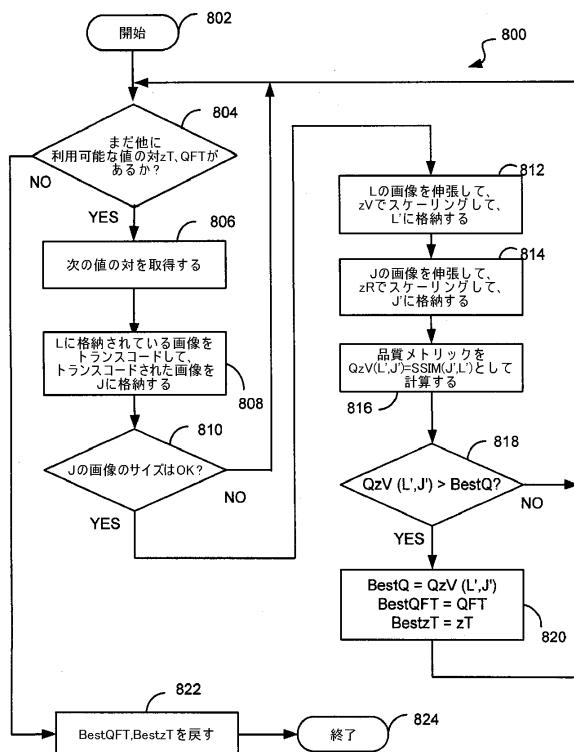
【図6】



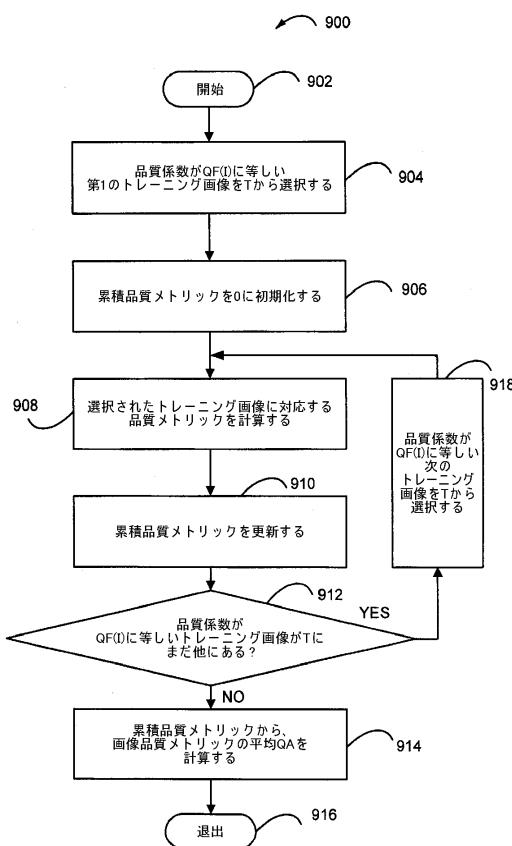
【図7】



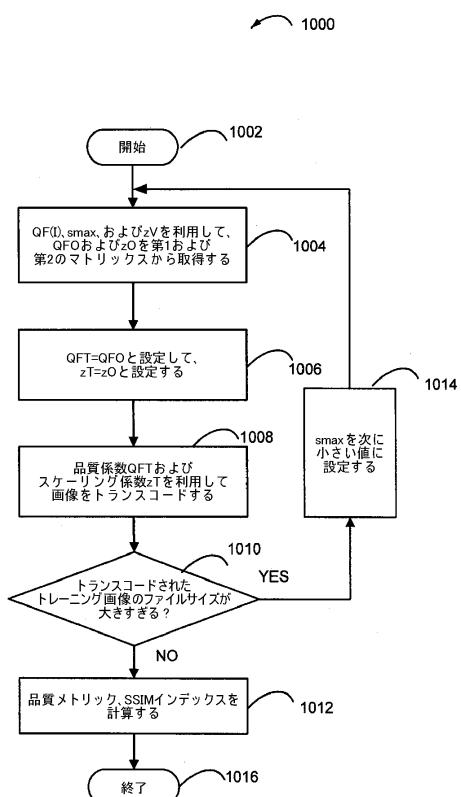
【図8】



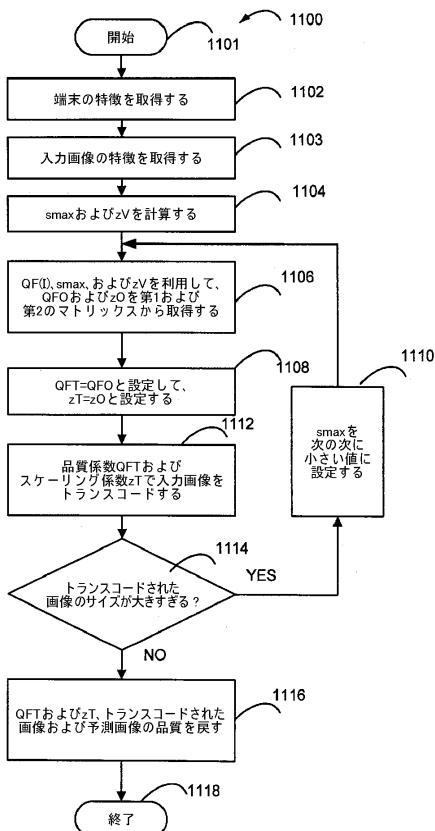
【図9】



【図10】



【図11】



【図12】

smax	10	20	30	40	50	60	70	80	90	100	zV, %
0.1	79.1	59.6	30.1	17.5	13.7	14.1	14.1	14.1	14.1	14.1	14.1
0.2	89.9	79.6	68.8	53.0	34.3	28.4	22.5	22.5	22.5	22.5	22.5
0.3	90.0	88.7	79.8	69.7	59.1	45.8	39.1	32.8	32.8	32.8	32.8
0.4	90.0	89.9	80.0	79.8	69.8	59.6	51.5	48.8	41.4	39.9	39.9
0.5	90.0	90.0	86.5	80.0	79.4	69.8	62.3	59.7	49.5	40.0	40.0
0.6	90.0	90.0	89.2	80.0	80.0	76.1	69.7	69.7	59.7	49.7	49.7
0.7	90.0	90.0	89.9	83.8	80.0	79.9	79.8	70.0	69.3	59.7	59.7
0.8	90.0	90.0	90.0	86.9	80.0	80.0	80.0	78.6	70.0	60.0	60.0
0.9	90.0	90.0	90.0	88.8	82.1	80.0	80.0	79.9	70.0	60.0	60.0
1.0	90.0	90.0	90.0	89.5	84.0	80.0	80.0	80.0	79.1	80.0	80.0

Figure 12

【図13】

smax	10	20	30	40	50	60	70	80	90	100	zV, %
0.1	10	20	28	32	33	33	33	33	33	33	33
0.2	10	20	30	40	48	50	52	52	52	52	52
0.3	10	20	30	40	50	59	60	63	63	67	67
0.4	10	20	30	40	50	60	69	70	71	76	76
0.5	10	20	30	40	50	60	70	70	79	88	88
0.6	10	20	30	40	50	60	70	70	80	90	90
0.7	10	20	30	40	50	60	60	79	80	90	90
0.8	10	20	30	40	50	60	69	71	80	90	90
0.9	10	20	30	40	50	60	70	70	80	100	100
1.0	10	20	30	40	50	60	70	79	80	100	100

Figure 13

【図14】

zV, %											
smax	10	20	30	40	50	60	70	80	90	100	zV, %
0.1	0.23	0.33	0.35	0.36	0.38	0.37	0.37	0.37	0.37	0.37	0.37
0.2	0.25	0.41	0.52	0.56	0.56	0.60	0.56	0.57	0.57	0.57	0.57
0.3	0.25	0.44	0.55	0.63	0.67	0.68	0.69	0.68	0.68	0.67	0.67
0.4	0.25	0.46	0.58	0.66	0.70	0.73	0.73	0.75	0.74	0.75	0.75
0.5	0.25	0.46	0.58	0.66	0.74	0.76	0.77	0.79	0.79	0.80	0.80
0.6	0.25	0.46	0.61	0.67	0.75	0.78	0.80	0.81	0.82	0.84	0.84
0.7	0.25	0.46	0.62	0.68	0.75	0.80	0.82	0.84	0.84	0.86	0.86
0.8	0.25	0.46	0.62	0.70	0.75	0.81	0.84	0.85	0.86	0.89	0.89
0.9	0.25	0.46	0.62	0.72	0.76	0.81	0.85	0.86	0.87	0.94	0.94
1.0	0.25	0.46	0.62	0.73	0.77	0.81	0.85	0.88	0.88	0.99	0.99

Figure 14

フロントページの続き

(72)発明者 ピジョン、スティーブン

カナダ国ケベック州 アーシュ3セー・1カ3 モントリオール, リュ・ノートル・ダム・ウエスト
ト 1100 エコール・ドゥ・テクノロジー・スュペリュール内

審査官 岩井 健二

(56)参考文献 特許第5176175(JP, B2)

特許第4994500(JP, B2)

特開2002-374298(JP, A)

特開2001-112006(JP, A)

特開2000-333163(JP, A)

国際公開第2006/126260(WO, A1)

国際公開第2006/098226(WO, A1)

国際公開第01/58096(WO, A1)

Yong Wang et al, Utility-Based Video Adaptation for Universal Multimedia Access (UMA)
and Content-Based Utility Function Prediction for Real-Time Video Transcoding, IEEE TR
ANSACTIONS ON MULTIMEDIA, IEEE, 2007年 2月, VOL. 9, NO. 2, pp. 213-220

(58)調査した分野(Int.Cl., DB名)

H04N 19/00 - 19/98