

100/PT62251(PL)5 案に対して出された意見とその処理

2002-01-16

ソース	意見	主査(PL)見解
W-01 渡辺馨 2002-01-11	(1) 5.1 節について 以下の(5)と関係しますが、結果の報告例を示す目的であれば、具体的なシーケンス名をこの節で示す必要はあるのでしょうか。結果の比較のためには同一のシーケンスを用いる必要があると思いますので、例えば、Table 1 のシーケンスについては、使用可能なシーケンスの例として Annex にまわされてはいかがでしょうか。	主観評価法と相関をもつものが意味のある客観評価法といえます。国際レベルで主観評価値とともに広く利用可能な参照ビデオを使っています。これを客観化するのを客観評価法としています。つまり参照ビデオをネットワーク等に流し受信した時、それを主観評価するのではなく、同一の客観化を適用し客観評価で置き換えようとしています。VQEG/CRC ビデオシーケンスと対応する主観評価値がオーソライズされていない段階では、「客観化」手続きそのものが、使用したビデオシーケンスに依存すると考えられますので、先頭に条件として挙げました。 Annex にまわすことについては異存ありません。具体的な数値例では、本文中でそれを参照することとします。
W-02 渡辺馨 2002-01-11	(2) 5.4 節の追加について 新たに追加された 5.4 節の 5.4.2 の(1)式の定義は、5.6 節(以前の 5.3 節)の 5.6.2 の(9)式と等価のように見受けられます。もし、そうであれば、5.4 節の内容は、5.6 節の PSNR に含めることでいかがでしょうか？	5.4 と 5.5 はビデオ対の誤差そのものを扱い、5.6 は誤差の逆数の対数を計算するという観点から分類しました。この観点から、(1)式と 5.6 の式(PSNR)とは異なります。しかし、前出の式を参照することによって、2 回同じような式が出ることを避けることとし、ご意見に沿いたいと存じます。
W-03 渡辺馨 2002-01-11	(3)5.5 節の追加について wavelet を用いた測定法は視覚特性等を考慮した客観評価尺度の一提案と考えます。 ビデオの主観値に関連した測定法は、ご存知のように VQEG で多数提案されていますので、wavelet 測定手法だけを取り上げることにについては議論があると思います。 例えば、"end-to-end objective assessment of video quality"については、本文中では、ITU や VQEG での審議されている内容や、提案されている方式の名称について記載し、wavelet 測定手法については一例として Annex に記載すること、などはいかがでしょうか。 また、章立てについても、考慮する要素や処理がだんだん複雑になる順で記載したほうが読者にとって読みやすいと考えますので、objective assessment については PSNR の後にしたほうが適当と考えます。	Wavelet 係数の差は視覚的空間周波数特性にある程度関係した量となりうると想像されます。本 IEC/TR ではそこまでは言及せず、あくまでも物理測定に立脚した客観評価法としたいのです。つまり、VQEG での主観評価結果を勘案して、それと可能な限り関係した(相関のある)測定データの処理法について調査することが必要と考えています。 振幅空間であれ、周波数空間であれ、物理的に測定してデータ化可能であり、それを誰がどこで実施しても同じ結果を与えるのを客観的方法としています。この観点からは、wavelet 変換を明確に表に出す必要はなく、また、ご指摘のように読みやすいように書き改めます。
W-04 渡辺馨 2002-01-11	(4)5.6 節について PSNR については、objective assessment と呼ぶほどには、視覚特性等を反映していない(色については、等色空間で計算した場合、一部含んでいるとはいえませんが)ため、節のタイトルは、"end-to-end objective assessment of video quality using PSNR"ではなく、これまで通り"PSNR"として、"end-to-end colour reproduction"の後にしたらどうでしょうか。 通常の測定では、色空間上の 3 次元の PSNR よりも、輝度や色差それぞれの 1 次元の PSNR がよく用いられます。したがって、ここでの記載も、Y や L などの 1 次元の PSNR の定義をしたうえで、3 次元 PSNR について定義されてはどうか。	この IEC/TR では客観評価法(object assessment)を W-03 への回答に書いたように理解しています。要は、振幅空間、周波数空間を問わず、また、振幅空間ではいかなる 3D 空間であるかを問わず、冗長測定によって得られるデータによっていかに主観評価を説明できるかが重要と考えています。この観点から、いかなるデータが意味あるデータであり、それをいかにデータ処理するかが客観評価法であり、その良さはどの程度主観評価を説明できているかで評価されるべきでしょう。 検討の結果、単純な PSNR は本 IEC/ TR には含めない方向で考えています。

ソース	意見	主 査 (PL) 見 解
W-05 渡辺馨 2002-01-11	(5)評価結果の報告形式について "Form of reporting assessment results"は、測定結果そのものよりは、結果の提示方法をどうすべきか簡潔にわかることが、本 TR の趣旨と考えます。したがって、複数の映像を用いてその測定結果を示すよりは、一つの映像を例として提示方法を示したほうが適当ではないでしょうか。以前の WD では、PSNR の結果について、用いた映像の絵柄（写真）とフレームごとの結果を示すグラフとシーケンスの平均を示すことを記載していましたが、その形式をあるシーケンスについて示されることが適当と考えます。	同意します。 書き改めます。 なお、先の IEC/PT62251 国際会議で、写真はやめた方が無難(Henk Kolk)との意見が出されたので、掲載しない方向で考えています。
W-06 渡辺馨 2002-01-11	(6)Frame rate ,ビット解像度、遅延の項について 最初の WD には、上記の項目が記載されていたと記憶しているのですが、最新版ではこれらの項目が削除されています。どのような経緯で削除することになったか再度お知らせ頂ければ幸いです。特に 6 章のオーディオの項には、これらの項目がありますので、できれば統一したほうがよいかと思います。	http://www.map.tu.chiba-u.ac.jp/IEC/PT62251 にある 2001-09-14 付の文書 100/PT62251(PL)3 には含まれていません。これは国内でご確認いただき、国際 PT62251 の意見聴取用です。これを改定したのが 2001-09-24 付の 100/PT62251(PL)3A で、ITU-R WP6Q に提出し、2001-10-19 の PT62251 国際会議で審議したものです。したがって、ご指摘の項目は公表された文書には最初からなかったこととなります。 (期限は 2001-12 となっており、大分遅延しておりますが) 具体的な測定法のご提案があれば IEC/TR に含めさせていただきます。
W-07 渡辺馨 2002-01-12	(1)対象とする映像信号について AV 通信システムで用いられる映像信号は、各種想定されると考えますが、コンピュータなどで扱う sRGB 信号と放送などで扱う TV 映像信号では、次のような違いがあります。 sRGB とハイビジョンの BT.709 の白色と 3 原色は、同じです。 <ul style="list-style-type: none"> ● ガンマ特性 ● レベル定義 (放送(8bit)では、R,G,B,Y: 16 (黒) ~ 235 (白 100%)、220 レベル。CB,CR:16 ~ 240、225 レベル。 ● 信号帯域 (放送のスタジオ信号は、4:2:2 の場合、CB,CR の帯域は Y の 1/2。MPEG 等の圧縮では 4:2:0 の場合、CB,CR は水平、垂直方向で 1/2 で処理。) このほか、100%レベル以上や負の信号も認める、などの違いも考えられます。 したがって、どのような映像信号を対象とするかで規定が変わってくるところがあり、本規定の対象とする範囲を明記する必要があります。 例えば、5.2 および 5.3 の規定は sRGB を想定していますので、それを明記し、その他の信号の場合、sRGB へ変換を行うか、別途規定が必要と記載することが適当と考えます。 また、5.1 で挙げられている標準画像は TV 信号ですので、その信号を用いて測定する場合、sRGB 信号との相違に配慮する必要がありますので、信号源として sRGB 信号の映像を例とされることが望ましいと思います。	先ず、IEC 62251 は Technical Report (TR, 技術情報)であり、いわゆる「規定」(normative)ではありません。ISO/IEC 共通の定義通りの文書です。 調査対象の範囲を明確にすることには同意します。典型的には、PC に用意された AVI 形式のファイル (R,G,B ピクセルの列)です。それを Network encoder が適当な形式に streaming encode して受信される。受信情報は表示装置に、最終的には R,G,B ピクセルとして得られる。これを end-to-end と称しています (Figure 2 参照)。 IEC/TR に盛り込むビデオ品質客観評価法の能力の確認・デモに使ったのは、対応する主観評価値が既知の VQEG/CRC のビデオシーケンスです。しかし、4:2:2 YUV 形式のファイルをそのまま使ったのではなく、これを ITU-R BT.601-5 に基づいて、sRGB データに変換し、1 ピクセルごと疎抜き、さらの周りの不要部分を捨て、フレームサイズを 320x240 ピクセルとして使っています。 このようにフレーム形式を変更した場合に、VQEG が与える客観評価値がそのまま有効なものとして使えるが否かの疑問が生じますが、近似的にそのまま使っています。 Reference video と Degraded video の sRGB 以外の信号形態については本 IEC/TR では扱わないこととします。 この辺を 5.1 に明確に示します。

ソース	意見	主査(PL)見解
<p>W-08 渡辺馨 2002-01-12</p>	<p>(2)色空間上での PSNR 定義について</p> <p>5.6 に記載されている PSNR のうち、私たちが良く用いているのは、R,G,B あるいは Y,CB,CR それぞれについて(7)式で計算したものです。</p> <p>(8)~(10)式のように色空間上で MSE を計算する場合、視覚的な特性を考慮する要素が生じ、例えば、CIELAB の各値で重み付けするなど、学会でもまだ議論がされていると聞いております。</p> <p>特に、上で述べたように、TV 信号で Y と C でダイナミックレンジや信号帯域幅が値が違う場合には、(8)~(10)式で表現することは難しいと考えます。</p> <p>AV 機器の測定においては、色空間上で PSNR を計算しなくとも、各映像信号について PSNR を求めることでも有用性はあると考えますので、色空間上での PSNR は今後の課題として扱われてはいかがでしょう。</p> <p>以上のことから、本 TR は想定する映像信号を sRGB とし、その他の信号については、sRGB に変換することを前提することとしてはいかがでしょう。</p> <p>ただし、信号の帯域幅など扱いに注意を要することも注釈として必要と思います。</p> <p>さらに、5.6 における色空間上での PSNR や、先日コメントしました 5.4、5.5 の内容については、学会等での議論を考慮いただいて、今後の課題とされるか Annex にまわすなどの無難な表現にしていただければ幸いです。</p> <p>細かい表現については十分見直すことはできておりませんが、以上、ご検討いただければ幸いです。</p>	<p>PSNR は peak-signal-to-noise ratio ですから、その定義を広くとらえれば、両者(peak-signal と noise)が定義できれば、いかなる 3D 空間でも客観値として計算できます。要は、その値がいかなる意味を持つか、つまりビデオ品質の主観評価値と関連付けられるものであるか否かでしょう。これについては、理論的・生理学的な検討も必要でしょうが、本 IEC/TR を準備するにあたっては、既知の主観評価値をベースに、実証的・実験的・発見的手法で、適当と考えられるデータを抽出しました。データが適当か否かはそのデータの処理方法と、最終結果がいかに基礎としている主観評価値を説明できるかで判断されます。</p> <p>検討の結果、W-04 への回答にも示したように、単純な PSNR は本 IEC/TR には含めない方向で考えています。</p> <p>これは、いただいたご意見に沿ったものとなるでしょう。</p>