Digital Image Analysis (CSL7320)

॥ त्वं ज्ञानमयो विज्ञानमयोऽसि ॥

Critical Review on "Towards a Simplified Perceptual Quality Metric for Watermarking Applications"

Submitted By:
Rati Kumari
(M20CS012)

Course Instructor:
Dr.Rajlaxmi Chouhan
Professor, Department of CSE
IIT Jodhpur

Introduction

New metrics are proposed to overcome the limitations of assessing distorted image quality in a specific frequency range. Previously used metrics were not efficient for watermarking applications and distortions measures. To provide better results a simplified objective quality metric is designed. This generic algorithm exploits the adapted Minkowski error pooling and CSF (contrast sensitivity function). A high correlation is observed between MOS (mean opinion score) and the proposed objective metric when comparing existing objective quality metrics. In the earlier time, the main focus was on robustness instead of invisibility. Commonly used techniques for watermarking are PSNR and SSIM. This paper highlights the limitations of these techniques when observed under different frequency ranges. The subjective experiment is conducted to assess watermarking algorithms' quality, where human observers are asked to analyze the image quality. A new objective quality metric is proposed considering the HSV features. Commonly used quality metrics have five steps.

1. Pre-processing step
2. CSF filtering
3. Masking
4. Error normalization
5. Error Pooling

For watermarking, Full Reference Quality Metric predicts the quality score by comparing both original and distorted images.

Keywords Used

MOS (MEAN OPINION SCORE)

It is used to measure the quality of the perceived image. MOS is obtained by taking input from different observers of the image who rate image quality within the range of [1 (bad) to 5 (excellent)].

MINKOWSKI ERROR POOLING

Minkowski error pooling is the final step of OQM, where *v(i)* represents the absolute difference at the i'th spatial location. N represents the number of pixels in an image.

$$Mink(P,R) = \sqrt[R]{\frac{1}{N}\sum_{i=1}^{N}|v_i|^P}$$

When P increases, the emphasis shifts to regions with high distortions. After fine-tuning the error pooling step, the results are better. When P=R, adjusting the parameters to get changes in the image's perceived quality is the flexibility.

SSIM (STRUCTURAL SIMILARITY INDEX MEASURE)

SSIM shows the resemblance between two images. Structural distortions are used to evaluate perceptual distortion based on the difference in luminance. This metric describes the image quality as it compares the image elements perceived by humans. SSIM considers the distortion due to luminance, contrast, and texture. Acceptable SSIM index is from -1 (maximum difference) to 1 (no difference). A larger value represents better quality.

PSNR (PEAK SIGNAL-TO-NOISE-RATIO)

PSNR is used to show the ratio of peak to noise. MSE (Mean Square Error) represents the maximum difference value. PSNR value is calculated using MSE, and a higher value represents a better value. For best quality, small MSE values and large PSNR values are considered. MSE and PSNR do not use the image elements perceived by humans.

VIF (VISUAL INFORMATION FIDELITY)

This is an assessment index for Input and Output of HVS. It is based on natural scene statistics and image information is obtained from HVS.

Summary of Proposed Metric

Objective metrics are designed considering two important concepts like contrast sensitivity and error pooling. Contrast sensitivity means a person's ability to distinguish between the foreground and background by observing the image's pixels. CSF represents the human's sensitivity to spatial frequency. Human Visual System is sensitive to the frequency response of images like contrast or luminescence level. When designing applications for watermarking, CSF is used to measure the distortions in different frequency ranges. 2D CSF filters use a Fourier spectrum to enhance the image quality where sensitivity is high. When watermarks are added in LF (Low Frequency), HF (High Frequency) and MF (Middle Frequency) then three distinct clusters corresponding to watermarks embedding are observed. Watermarked images with the same PSNR value can have varied perceived quality of the image if watermark frequency is changed. OQMs behaviors are directly proportional to the distortions in the frequency range. As decomposition of the perceptual channel is computationally expensive, it is not considered in the proposed metric. The non-linearity function does not impact the MOSp value, so the resulting value depends on the viewing screen.

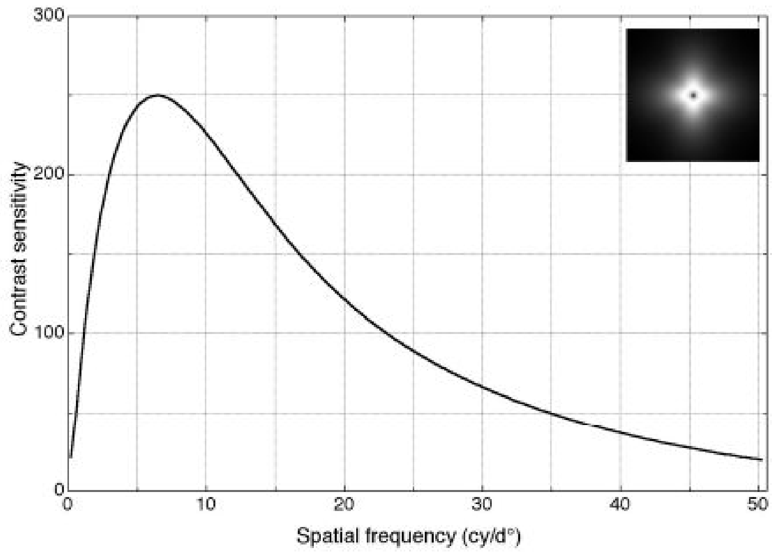

Fig.1. Spatial frequency vs Contrast sensitivity.

HVS based OQM has five steps and has more computation. So, the proposed metric (Fig.3) is restricted to a frequency weight and error values. The proposed metric has 2D-Fourier transformation on both original and distorted images. Then the outcome is weighted by 2D-CSF followed by inverse Fourier transform on both resulting spectrums.

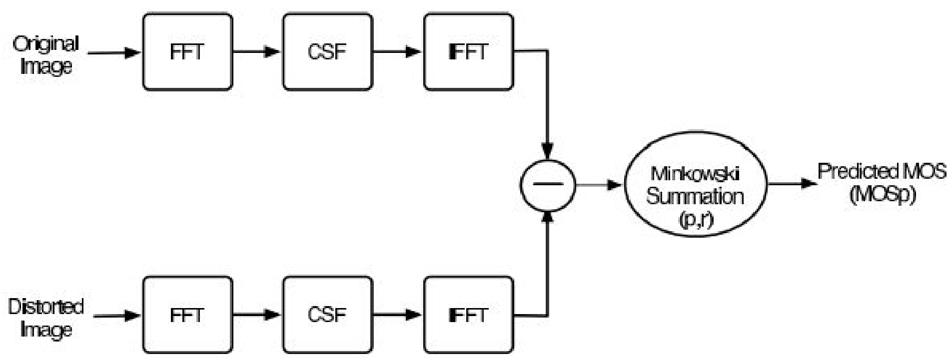

Fig.3. Block diagram of the proposed "CPA" Metric.

The resultant spectrum is computed and a weighted error map is obtained followed by Minkowski summation of error map which gives MOSp value. Fine tuning on the error value gives better results.

Experiment

The performance of CPA is analyzed on four different subjective datasets with three datasets from digital watermarking transformed spaces and the fourth one from coding datasets for comparison.

The MOS values are normalized according to standards set by ITU (International Telecommunication Union). The images are placed at a distance of six times the display height. Impairments were on a scale of 5 categories (1. "Very Annoying", 2. "Annoying", 3. "Slightly annoying", 4, "Perceptible, but not annoying", 5. "Imperceptible"). The database used for testing are:

Database1: The database has images marked using FFT-domain watermarking techniques. The observers are seven in number. Input are Gray level images and are marked independently for different frequency ranges and perceptual subbands.

Database2: The database has perceived distortion of 120 images. It has Seventeen observers. In this, grayscale images are used, and "broken arrows" embedding techniques generate different embedding strengths.

Database3: The database has grayscale images as input with fourteen observers. The watermarks are multiplicatively embedded onto the subbands. The dataset has perceived 120 distorted images.

Database4: The database has nineteen observers. The dataset has perceived 120 distorted images. The database has coding distortions of type (JPEG2000, JPEG, and LAR).

The performance was analyzed after comparing with twelve metrics from the "metrix_mux" package§ (MSE, SSIM, PSNR, MSSIM, VIF, VIFP, UQI,VSNR, IFC, NQM, SNR, WSNR).

CRITIQUE

The performance is evaluated based on VQEG (Video Quality Experts Group) metrics. The fitting function is used to match objective scores in the subjective range according to VQEG.

Metric	Pearson Correlation				Spearman Correlation			
	DB1	DB2	DB3	DB4	DB1	DB2	DB3	DB4
MSE	0.435	0.880	0.877	0.793	0.697	0.929	0.873	0.817
PSNR	0.738	0.953	0.918	0.836	0.697	0.929	0.873	0.817
SSIM	0.516	0.846	0.927	0.899	0.640	0.878	0.926	0.885
MSSIM	0.761	0.854	0.909	0.914	0.865	0.896	0.919	0.919
VSNR	0.907	0.834	0.746	0.893	0.907	0.827	0.755	0.862
VIF	0.873	0.944	0.895	0.947	0.862	0.934	0.890	0.917
VIFP	0.723	0.939	0.914	0.938	0.695	0.931	0.916	0.896
UQI	0.574	0.904	0.953	0.917	0.600	0.907	0.951	0.876
IFC	0.846	0.884	0.933	0.937	0.837	0.879	0.912	0.914
NQM	0.382	0.913	0.794	0.943	0.381	0.905	0.784	0.920
WSNR	0.771	0.887	0.898	0.927	0.793	0.871	0.887	0.910
SNR	0.736	0.892	0.892	0.783	0.689	0.883	0.886	0.774
WPSNR	0.756	0.952	0.939	0.940	0.707	0.949	0.928	0.926
C4	0.822	0.931	0.935	0.955	0.853	0.932	0.925	0.929
KOMP	0.674	0.856	0.915	0.921	0.764	0.888	0.907	0.911
CPA	0.913	0.974	0.951	0.946	0.920	0.969	0.948	0.918

Table.1. Pearson and Spearman correlations

The Minkowski parameters P and Q are examined in the range [1,12] with the step of 0.5. Best results are observed when P=5, R=10. The CPA metrics present the best result on four databases with a correlation above 0.9. Tables 1 and 2 represent the performance metrics. The results of the metrics are analyzed for RMSE / Outlier Ratio and Pearson / Spearman Correlations.

Metric	RMSE				Outlier Ratio			
	DB1	DB2	DB3	DB4	DB1	DB2	DB3	DB4
MSE	1.124	0.578	0.886	0.821	0.581	0.233	0.575	0.483
PSNR	0.840	0.366	0.491	0.739	0.400	0.092	0.242	0.475
SSIM	1.148	0.643	0.475	0.590	0.614	0.333	0.167	0.367
MSSIM	0.809	0.669	0.524	0.587	0.419	0.350	0.192	0.400
VSNR	0.525	0.665	0.826	0.606	0.210	0.342	0.433	0.342
VIF	0.638	0.398	0.552	0.434	0.319	0.125	0.267	0.242
VIFP	0.860	0.416	0.502	0.466	0.448	0.133	0.192	0.317
UQI	1.020	0.514	0.393	0.539	0.510	0.200	0.125	0.308
IFC	0.664	0.563	0.446	0.470	0.310	0.217	0.158	0.250
NQM	1.151	0.493	0.754	0.447	0.548	0.167	0.333	0.250
WSNR	0.793	0.556	0.546	0.506	0.386	0.233	0.242	0.275
SNR	0.843	0.545	0.561	0.838	0.410	0.225	0.267	0.467
WPSNR	0.815	0.37	0.426	0.462	0.390	0.067	0.175	0.267
C4	0.717	0.442	0.440	0.400	0.352	0.142	0.175	0.208
KOMP	0.946	0.775	0.516	0.556	0.510	0.467	0.225	0.325
CPA	0.508	0.273	0.382	0.439	0.200	0.025	0.108	0.267

Table.2. RMSE and Outlier ratio

The distribution of points is linear. The proposed metrics "CPA", measure the difference in both original and distorted images. The straight-line distribution shows the narrowest distribution points obtained by CPA metrics. The proposed metric shows better performance on four databases than other metrics and best performance on database 1. When SSIM and PSNR are checked for watermark with different frequency modulation, the results were not good. Due to different frequencies embedding discrepancy was observed on database1. C4 metric has given average results on all databases except the fourth one, which offers the best impact on coding distortions.

Inferences

The proposed simple Objective Quality Metrics uses the concepts of Human Visual System. When it is compared with fifteen state of the art metrics on datasets with varying distortions, the proposed metrics was more efficient. CPA metric gives the best performance when the watermark is modulated with a different frequency range. MOS vs MOSp comparison showed that it could be used to get better objective quality in watermarking techniques.

References

[1] Carosi, Maurizio & Pankajakshan, Vinod & Autrusseau, Florent. (2010). Towards a Simplified Perceptual Quality Metric for Watermarking Applications. Proc SPIE. 7542. 10.1117/12.838627.

[2] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, “Image quality assessment: From error visibility to structural similarity,” IEEE Transactions on Image Processing 13(4), pp. 600–612, 2004.

[3] S. Voloshynovskiy, S. Pereira, A. Herrigel, N. Baumgartner, and T. Pun, “Generalized watermarking attack based on watermark estimation and perceptual remodulation,” in SPIE Security and Watermarking of Multimedia Contents, 3971, pp. 358–370, 2000.

[4] T. Furon and P. Bas, “Broken arrows,” EURASIP Journal on Information Security 2008, pp. 1–13, 2008.

[5] D. Barba and P. Le Callet, “A robust quality metric for color image quality assessment,” in IEEE International Conference on Image Processing, pp. 437–440, 2003.