Performance Analysis of Structure Similarity Algorithm for the Recognition of Handwritten Urdu Characters

Ms. Zareen Yakub Khan[1], *Dr. P. K. Butey*[2], and *Dr. G.S. Katkar* [3]
[1]*Research Scholar, Kamla Nehru College, Nagpur*
[2]*Assistant Professor+, Kamla Nehru College, Nagpur*
[3]*Assistant Professor, Taywade College, Mahadula-Koradi*

***Abstract*: In this paper we analyse the performance of Structure Similarity algorithm for the recognition of printed cursive english alphabets, ssim is a structural method suitable for printed characters but not efficient for cursive handwriiten characters.**

Published in: Journal of the Asiatic Society of Mumbai ISSN: 0972-0766 UGC Care Group 1 Journal.

STRUCTURE SIMILARITY INDEX ALGORITHM

Perceptual image quality metrics have explicitly accounted for human visual system (HVS) sensitivity to sub-band noise by estimating thresholds above which distortion is just-noticeable. Objective methods for assessing perceptual image quality traditionally attempted to quantify the visibility of errors (differences) between a distorted image and a reference image using a variety of known properties of the human visual system. A recently proposed class of quality metrics, known as structural similarity (SSIM), models perception implicitly by taking into account the fact that the HVS is adapted for extracting structural information (relative spatial covariance) from images and specific SSIM implemented both in the image space and the wavelet domain.

The motivation behind the structural similarity approach for measuring image quality is that the HVS has evolved to do visual pattern recognition in order to be able to extract the structure or connectedness of natural images. Based on this observation, it makes sense that a useful perceptual quality metric would emphasize the structure of scenes over the lighting effects. If PSNR (Peak Signal to Noise Ratio) value decreases quality of Image Increases and Vice versa with respect to the different image enhancement technique will give the different values for PSNR.

The structural similarity (SSIM) index is a method for measuring the similarity between two images. The SSIM index is a full reference metric, in other words, the measuring of image quality based on an initial uncompressed or distortion-free image as reference. SSIM is designed to improve on traditional methods like peak signal-to-noise ratio (PSNR) and mean squared error (MSE), which have proven to be inconsistent with human eye perception. The difference with respect to other techniques mentioned previously such as MSE or PSNR is that these approaches estimate perceived errors; on the other hand, SSIM considers image degradation as perceived change in structural information. SSIM is commonly used for testing the quality of various lossy images. The SSIM index is a decimal value between 0 and 1. A value of 0 would mean zero correlation with original image, and 1 means the exact same image, so through this index, image comparison methods can be effectively compared. It is defined as:

$$SSIM(x,y) = [1(x,y)]^{\alpha}\, c[(x,y)]^{\beta}\, s[(x,y)]^{\gamma}$$

----------(5.1)

Here *(x, y)* are two images. *1 (x, y)* is the luminance comparison. *c (x, y)* is contrast comparison and s*(x,y)* is the structural comparison between two images and $\alpha>0$, $\beta>0$ and $\gamma>0$ are used to adjust the importance of three parameters. In practice, one usually requires a single overall quality measure of the entire image, so mean SSIM (MSSIM) index is used to evaluate the overall image quality. The difference with respect to other techniques mentioned previously such as MSE or PSNR is that these approaches estimate perceived errors; on the other hand, SSIM considers image degradation as perceived change in structural information. Structural information is the idea that the pixels have strong inter-dependencies especially when

they are spatially close. These dependencies carry important information about the structure of the objects in the visual scene. The SSIM metric is calculated on various windows of an image.

Structural information is the idea that the pixels have strong inter-dependencies especially when they are spatially close. These dependencies carry important information about the structure of the objects in the visual scene. The SSIM metric is calculated on various windows of an image. The measure between two windows x and y of common size $N \times N$ can represent as:

$$SSIM(x,y) = \frac{2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_{2)}} \quad (5.2)$$

With μ_x is the average of x, μ_y is the average of y; σ_x^2 the variance of x; σ_y^2 the variance of y; σ_{xy} the covariance of x and y and $c_1=(k_1 L)^2$, $c_2=(k_2 L)^2$ two variables to stabilize the division with weak denominator.

In order to evaluate the image quality this formula is applied only on *LUMA*. *LUMA* represents the brightness in an image (the "black-and-white" or achromatic portion of the image). The resultant SSIM index is a decimal value between -1 and 1, and value 1 is only reachable in the case of two identical sets of data. Typically, it is calculated on window sizes of 8×8. The window can be displaced pixel-by-pixel on the image but the authors propose to use only a subgroup of the possible windows to reduce the complexity of the calculation. Structural dissimilarity (DSSIM) is a distance metric derived from SSIM (though the triangle inequality is not necessarily satisfied).

$$DSSIM(x,y) = \frac{1 - SSIM(x,y)}{2} \quad (5.3)$$

The peak signal-to-noise ratio, often abbreviated PSNR, is an engineering term for the ratio between the maximum possible power of a signal and the (or codec type) and same content. It is most easily defined via the mean squared error (MSE) which for two m × n monochrome images I and K where one of the images is considered a noisy approximation of the other is defined as MSSIM i.e. Measurement of structure similarity algorithm used for image structure analysis [79]. Thus the mean square error can define as [80]:

$$MSE = \frac{1}{mn}\sum_{i=0}^{m-1}[I(i,j) - K(i,j)^2 \quad (5.4)$$

Researcher has utilized this PSNR to judge the improved image quality with the help of MSE as follows:

$$PSNR = 10\log_{10}(\frac{MAX_1^2}{MSE}) \quad -- \quad (5.5)$$

$$= 20\log_{10}(\frac{MAX_1}{\sqrt{MSE}}) \quad -------(5.6)$$

Here, MAX_I is the maximum possible pixel value of the image. When the pixels are represented using 8 bits per sample, this is 255. More generally, when samples are represented using linear PCM with B bits per sample, MAX_I is $2B-1$. For color images with three RGB values per pixel, the definition of PSNR is the same except the *MSE* is the sum over all squared value differences divided by image size and by three [81]. Typical values for the PSNR in noisy image and video compression are between 30 and 50 dB, where higher is better. When the two images are identical the MSE will be equal to zero, resulting in an infinite PSNR.

Introduction to SSIM

The predecessor of SSIM was called Universal Quality Index (UQI), or Wang–Bovik Index, and was developed by Zhou Wang and Al Bovik in 2001. This evolved, through their collaboration with Hamid Sheikh and Eero Simoncelli, into the current version of SSIM, which was published in April 2004 in the IEEE Transactions on Image Processing.[1]. In addition to defining the SSIM quality index, the paper provides a general context for developing and evaluating perceptual quality measures, including connections to human visual neurobiology and perception, and direct validation of the index against human subject ratings.

SSIM was rapidly adopted by the image processing community, in part because the March 2000 FRTV Phase I report by the Video Quality Experts Group had concluded that nine previously proposed models for perceptual quality were ineffective.[2]

The structural similarity (SSIM) index is a method for predicting the perceived quality of digital television and cinematic pictures, as well as other kinds of digital images and videos. The first version of the model was developed in the Laboratory for Image and Video Engineering (LIVE) at The University of Texas at Austin and further developed jointly with the Laboratory for Computational Vision (LCV) at New York University.

SSIM is used for measuring the similarity between two images. The SSIM index is a full reference metric; in other words, the measurement or prediction of image quality is based on an initial uncompressed or distortion-free image as reference. SSIM is designed to improve on traditional methods such as peak signal-to-noise ratio (PSNR) [3][4] and mean squared error (MSE).

By the Help of MATLAB this is the formula from where we can judge whether the image quality is improved or not after process out MSSIM (measurement of Structure similarity) must be near to one and

```
PSNR=20*log(255/std2(abs(j-i(:,:,1))));      (5.7)
```

Mathematical Expressions

Peak signal-to-noise ratio (PSNR) and Mean Squared Error (MSE)

The peak signal-to-noise ratio, often abbreviated PSNR, is an engineering term for the ratio between the maximum possible power of a signal and the (or codec type) and same content [5] [6]. It is most easily defined via the mean squared error (MSE) which for two m×n monochrome images I and K where one of the images is considered a noisy approximation of the other is defined as [5],

MSSIM –is a Measurement of structure similarity algorithm [7] used for image structure analysis describe in Annexure-A.

$$MSE=\frac{1}{mn}\sum_{i=0}^{m-1}[I(i,j)-K(i,j)^2 \quad (5.8)$$

Researcher has utilized this PSNR to judge the improved image quality with the help of MSE the expression is carried out as follows.

The PSNR is can be defined for our experimental view as:

$$PSNR=10.\log_{10}(\frac{MAX_1^2}{MSE})=20\log_{10}(\frac{MAX_1}{\sqrt{MSE}})$$

Here, MAXI is the maximum possible pixel value of the image. When the pixels are represented using 8 bits per sample, this is 255. More generally, when samples are represented using linear PCM with B bits per sample, MAXI is $2B-1$. For color images with three RGB values per pixel, the definition of PSNR is the same except the MSE is the sum over all squared value differences divided by image size and by three [5].Typical values for the PSNR in lossy image and video compression are between 30 and 50 dB, where higher is better. When the two images are identical the MSE will be equal to zero, resulting in an infinite PSNR.

Standard Database Created for Hand written Urdu Characters

We have created standard database for Hand written Urdu Characters as shown in table 1. Each character image is a PNG image of size 50x50.

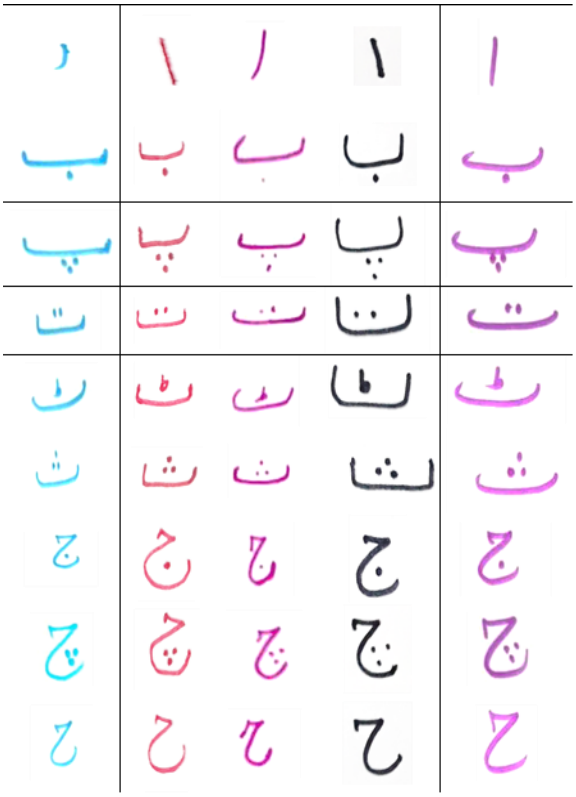

Table: Standard Database for Handwritten Urdu Characters

Experimental Work

Using MATLAB image processing tool, the matrices of each Hand written Urdu Characters are created along with the network structure. Pull the binary input code from the matrix and interpret the binary output code. To test the effect of Hand written Urdu Characters on character recognition, Hand written Urdu Characters samples were scanned and it had been converted to the vectors. Each character image is converted to a MATLAB vector. Standard printed Urdu Characters were used to create the test set to determine the structure similarity between standard sample of Hand written Urdu Characters and corresponding printed Urdu Characters.

In the experiment we had taken standard images of the Hand written Urdu Characters with which the samples of printed Urdu Characters are compared. SSIM function generates the accuracy results as shown in the table 2.

Our goal is to provide the noise free image near to original image it has been checked by MSSIM (Measurement of Structure Similarity Algorithm) [6] as the MSSIM values increases the noise decrease so MSSIM and Noise value are running inversely proportional.

We applied our scheme on selected images out of 50 sample images obtained from different individuals. The data set was containing varieties of writing styles.

Following are the steps to follow in this experimental work:

1. Data collection in the form of printed characters.
2. Collection of data also includes some scanned data taken from old document.
3. Hence, noise was removed & tried to clarify at most pure letter, for this some image processing techniques are used like High pass filter and Gaussian filter.
4. After this, we applied SSIM algorithm to classify the letters from each other.
5. Most of the image processing tools and classification work has been carried out by using MATLAB 7.8 software.
6. Input image to the system is binary image of 50 X 50 sizes.

Observations

S. No.	Urdu Characters	Measured Structure Similarity				
		Sample 1	Sample 2	Sample 3	Sample 4	Sample 5
0		0.9759	0.9801	0.9770	0.9800	0.9756
1		0.9806	0.9806	0.9834	0.9856	0.9833
2		0.9766	0.9743	0.9812	0.9788	0.9765
3		0.9811	0.9813	0.9806	0.9845	0.9800
4		0.9771	0.9743	0.9722	0.9744	0.9712
5		0.9759	0.9759	0.9759	0.9759	0.9765
6		0.9723	0.9812	0.9745	0.9806	0.9788
7		0.9812	0.9806	0.9712	0.9766	0.9733

Table: Handwritten Urdu Characters Recognition Accuracy

As shown in the table 2 total 50 Hand written Urdu Characters were randomly selected by using simple random sampling method (lottery method), total 10 Hand written Urdu Characters were compared with corresponding printed Urdu Characters. Each 5 samples of Urdu Characters 0-9 are compared with standard Urdu Characters from standard data set.

DISCUSSION

In this experimental work, an attempt is made to apply measured structure similarity approach to off-line recognition of Hand written Urdu Characters. Maximum performance rate of SSIM was found to up to 98 percent which is most promising than other methods. These results suggest applying image enhancement techniques to remove the noise from the images of Hand written Urdu Characters and reapply structure similarity algorithm to check percentage of structure similarity between standard Urdu Characters and sample Hand written Urdu Characters.

RESULT ANALYSIS AND CONCLUSION

In this experimental work, an attempt is made to apply measured structure similarity approach to off-line recognition of printed Hand written Urdu Characters. Total 50 samples were tested using SSIM. The performance rate of SSIM was found to up to 98 percent which is most promising than other methods. Chaudhari [1997] reported success rate up to 95% using zonal & structural feature extraction techniques used for characters [8].

It is observed that if SSIM index value is below 0.5 then we need to apply image enhancement techniques. We had implemented following image enhancement techniques.

1. Removed noise and background from threshold image. We obtained highly acceptable results.
2. High pass filter to increase contrast of character and to illuminate background noise. We have used imadjust() - Results are quite acceptable.
3. Gaussian filter – Results are not acceptable for most of the noisy character image.

It is necessary to apply image enhancement techniques before classifying Hand written Urdu Characters.

SSIM is best suitable for the recognition of printed Hand written Urdu Characters. It is not efficient algorithm to recognize Handwritten Characters.

REFERENCE

[1] *Wang, Zhou; Bovik, A.C.; Sheikh, H.R.; Simoncelli, E.P. (2004-04-01). "Image quality assessment: from error visibility to structural similarity". IEEE Transactions on Image Processing. 13 (4): 600–612. doi:10.1109/TIP.2003.819861. ISSN 1057-7149.*

[2] *"VQEG FRTV Phase I". 2000.*

[3] A.C. Brooks, Z. Xiaonan, and T.N. Pappas, "Structural similarity quality metrics in a coding context: Exploring the Space of Realistic Distortions," IEEE trans. on Image Processing, vol. 17, no. 8, pp. 1261-1273, Aug. 2008

[4] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE trans. on Image Processing, vol. 13, pp. 600–612, Apr. 2004.

[5] Alan C. Brooks and Thrasyvoulos N. Pappas Structural similarity quality metrics in a coding context exploring the space of realistic distortions

[6] Z.Wang, E. P. Simoncelli, and A. C. Bovick, "Multi-scale structural similarity for image quality assessment,"37th IEEE Asilomar Conf. on Signals, Systems and Computers 37, 2003.

[7] J. Bigun and G. H. Granlund, "Optimal Orientation Detection of Linear Symmetry", First International Conference on Computer Vision, London, June 1987, IEEE Computer Society Press, Washington DC, pp. 433-438.

[8] Video Quality Experts Group (VQEG), http://www.its.bldrdoc.gov/vqeg.

[9] Z. Wang and E. P. Simoncelli, "Translation insensitive image similarity in complex wavelet domain," in IEEE Int. Conference on Acoustics, Speech, and Signal Processing, P. Philadelphia, ed., Proc. IEEE II,pp. 573–576, 2005.

[10] Chaudhari "Zonal & structural feature extraction techniques used for recognition of printed devnagari characters", 1997.