Objective Image Quality Assessment Based on Support Vector Regression

Manish Narwaria and Weisi Lin

Abstract—**Objective image quality estimation is useful in many visual processing systems, and is difficult to perform in line with the human perception. The challenge lies in formulating effective features and fusing them into a single number to predict the quality score. In this brief, we propose a new approach to address the problem, with the use of singular vectors out of singular value decomposition (SVD) as features for quantifying major structural information in images and then support vector regression (SVR) for automatic prediction of image quality. The feature selection with singular vectors is novel and general for gauging structural changes in images as a good representative of visual quality variations. The use of SVR exploits the advantages of machine learning with the ability to learn complex data patterns for an effective and generalized mapping of features into a desired score, in contrast with the oft-utilized feature pooling process in the existing image quality estimators; this is to overcome the difficulty of model parameter determination for such a system to emulate the related, complex human visual system (HVS) characteristics. Experiments conducted with three independent databases confirm the effectiveness of the proposed system in predicting image quality with better alignment with the HVS's perception than the relevant existing work. The tests with untrained distortions and databases further demonstrate the robustness of the system and the importance of the feature selection.**

Index Terms—**Image quality assessment, image structure, singular value decomposition (SVD), support vector regression (SVR).**

I. INTRODUCTION

Auto-assessing the quality of digital video/images plays a very crucial role in image and video processing, and many practical situations, such as process evaluation (benchmarking different algorithms), optimization (like for a video encoder), and monitoring (e.g., in transmission and manufacturing sites). In addition, how to evaluate picture quality plays a central role in shaping most (if not all) visual processing algorithms and systems as well as their implementation. Visual quality evaluation in line with the human perception is with high significance for research [1], [4], [5], [7]–[11], [13], [15] due to the fundamental nature and the challenge to mimic the human visual system's (HVS's) perception.

In this brief, our aim is to first select proper image features towards visual quality evaluation based upon the available relevant knowledge about the HVS and then use machine learning to model the complex process of feature pooling. Machine learning techniques have been successful in face detection, object categorization, content-based image retrieval, texture classification, handwriting recognition, image classification, and object detection, and we believe that they can be used to establish a relationship between the perceived quality and a set of image features by learning through examples. Obviously the success of such a machine-learning-based image quality predictor hinges on the image features and the understanding of the HVS can be helpful to select appropriate features which are representative of visual quality variations.

There are two important issues in objective image quality assessment: 1) extraction and representation of appropriate features; and 2) pooling of the features for the result to be consistent with the HVS's perception of visual quality.

The first issue is not easy and clear-cut since the HVS is too complex to be fully understood with the present psychophysical means. The well-established facts include that the HVS is sensitive to spatial frequency and structures [1], [2], [4], [5], and visual structure in images plays a major role in the recognition of image content [3]. Other factors (like luminance changes) play a relatively insignificant role. Therefore, there has been a growing interest to use image structure for picture quality evaluation. A well-cited metric-based upon structure has been proposed by Wang and Bovik: the structural similarity index (SSIM) [4]. The SSIM measures the luminance, contrast, and structure changes between two images to gauge the quality score. Several other quality assessment methods like [5] and [15] have been developed based upon various ways of edge contrast/sharpness evaluation. The basic idea behind these methods is to quantify the HVS' perception when viewing distorted images.

The second issue is also not straightforward since the contribution (i.e., weight) of each feature to the final quality score may be different and is very difficult to be determined. In other words, it is a challenge to combine the statistics that quantify distortions into a single score. Researchers have employed techniques like simple summation/averaging of errors [4], [13], Minkowski metric fusion [7], linear (i.e., weighted) combinations, etc. Such techniques implicitly make assumptions on the relative importance of distortion statistics, and there is lack of convincing ground for these assumptions. The task becomes even more difficult when the number of features is large, and a machine learning method is expected to be more effective and convincing than the existing pooling methods. There has been some early work in applying machine learning for visual quality evaluation. In [8] and [9], objective quality assessment of video using neural networks (NNs) has been reported while the use NNs has been demonstrated in [10] and [11] for image cases. However, in these approaches, effort has been not directed to feature selection that is very important in machine learning [6], and overall, machine learning in visual quality evaluation remains as a largely uninvestigated area. To the best of our knowledge, a generalized machine-learning-based image quality metric does not exist in the current literature.

In this brief, we first propose the use of singular vectors (instead of singular values as used in [13], for the reasons to be given in Section II) out of singular value decomposition (SVD) as the selected features for structure representation in images. The singular vectors denote clearer physical meaning of structural degradations in comparison with the existing metrics [4], [13], [15]. Although other choices of machine learning are possible, in this work, we use support vector regression (SVR) for feature pooling because of the relatively high dimensionality of the proposed feature vector out of SVD. The SVR is suitable and effective in handling high-dimensional data. Hence, we propose an SVR-based image quality prediction system with singular vectors for feature selection, to achieve better correlation with the subjective scores than the relevant existing quality metrics.

The rest of this brief is organized as follows. Section II presents the characterization of images by SVD and discusses the details of the proposed visual quality metric, while extensive experimental results are presented and discussed in Section III. Finally, Section IV gives the concluding remarks.

II. THE PROPOSED IMAGE QUALITY ASSESSMENT SCHEME

A. Singular Value Decomposition

SVD of an image $\mathbf{A} \in \mathbf{R}^{r \times c}$ can be written as

$$\mathbf{A} = \mathbf{U}\boldsymbol{\sigma}\mathbf{V}^{T} \tag{1}$$

where $\mathbf{U}, \mathbf{V}$ and $\boldsymbol{\sigma}$, respectively, denote the left singular vector matrix (with size $r \times r$), the right singular vector matrix (with size $c \times c$),
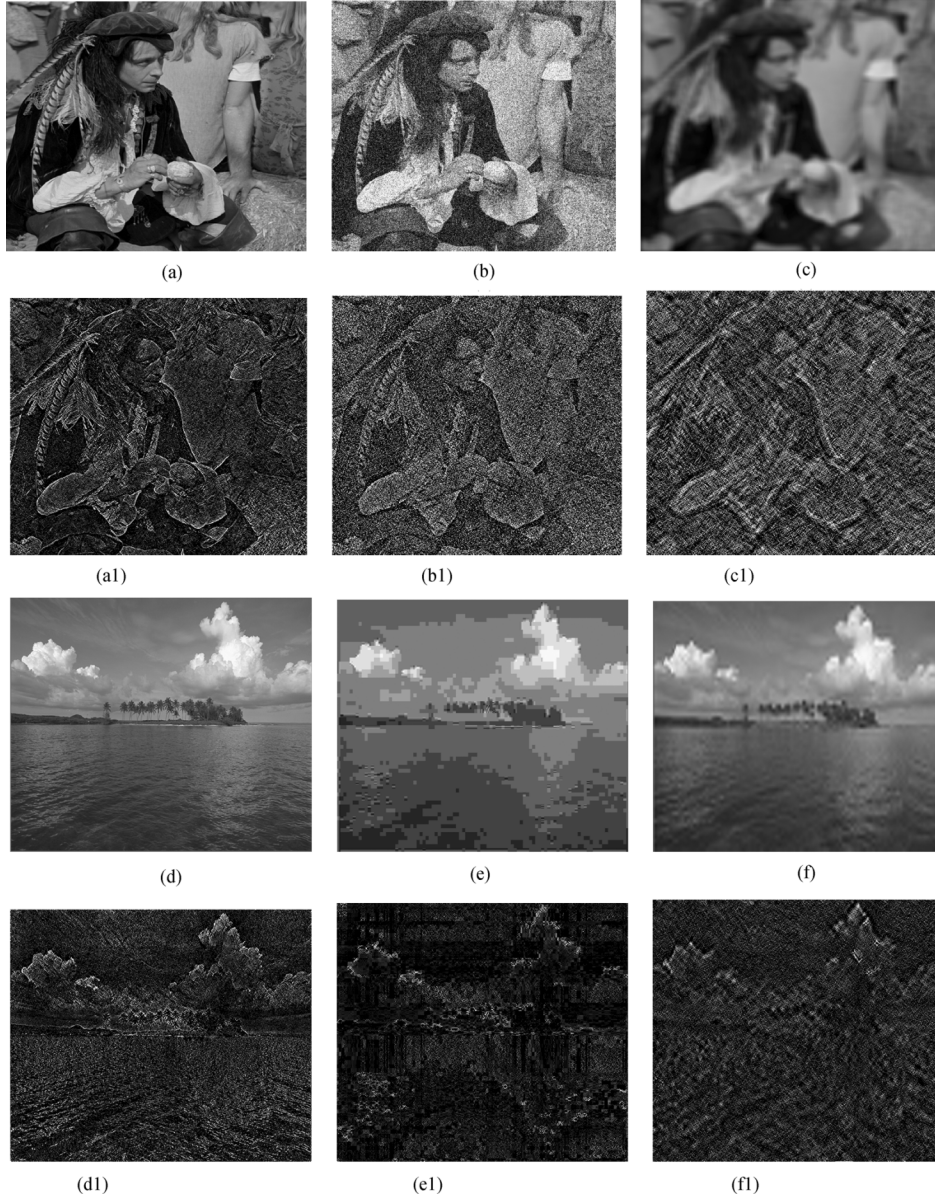
Fig. 1. Structure denoted by the singular vectors, i.e., $\mathbf{U}\mathbf{V}^T$ in images. (a) Original image. (b) Noisy image. (c) Blurred image. (a1) Original $\mathbf{U}\mathbf{V}^T$. (b1) Noisy $\mathbf{U}\mathbf{V}^T$. (c1) Blurred $\mathbf{U}\mathbf{V}^T$. (d) Original image. (e) JPEG image. (f) JP2K image. (d1) Original $\mathbf{U}\mathbf{V}^T$. (e1) JPEG $\mathbf{U}\mathbf{V}^T$. (f1) JP2K $\mathbf{U}\mathbf{V}^T$.

and the diagonal matrix (with size $r \times c$) of singular values. Let σ_j denote a singular value in $\boldsymbol{\sigma}$, $j = 1$ to z with $z = \min(r, c)$. Each σ_j corresponds to a part of the image energy (luminance) but not the image structure. $\mathbf{U}$ and $\mathbf{V}$ control the spatial distribution of the image energy. Fig. 1 shows the geometrical framework in images denoted by the product of $\mathbf{U}$ and $\mathbf{V}$ matrices (i.e., $\mathbf{U}\mathbf{V}^T$). It is well known that singular vectors are sensitive to perturbations [12]. It can be seen that adding distortion to an image results in the perturbation of the singular vectors. We can clearly see from Fig. 1 that the distortions modify the singular vectors and this leads to a distorted geometry of the image. Quantifying the structural distortions using singular vectors should therefore provide an effective basis for assessing image quality since the HVS is sensitive to structural changes [1], [2], [4], [5].

B. Feature Preparation

Different types of distortions (like JPEG artifacts, Blur, etc.) affect image quality in a largely similar fashion: by degrading image structure. By using $\mathbf{U}$ and $\mathbf{V}$, we exploit this common phenomenon in picture quality degradation. The central idea of the proposed scheme is to gauge structural modifications in images by measuring the changes in singular vectors. The singular vectors denote image structure which is more important for the HVS's perception of image quality, while singular values are less important with regards to image quality. In addition, for a same perturbation, the change in singular values $\boldsymbol{\sigma}$ is also reflected in that of singular vectors $\mathbf{U}$ and $\mathbf{V}$ (refer to [12]). In other words, singular vectors capture the major distorting factor (structural changes), and to certain extent, also reflect the minor one (luminance changes) due to the same perturbation. In this work, our main focus is on the extraction of the most crucial and meaningful information pertaining to image quality and adapt the SVR algorithm for effective image quality assessment. Thus, for the benefits of computational costs associated with lower dimension feature vector, we have not considered singular values in this work and used singular vectors as the feature basis for the task.

For image matrix $\mathbf{A}$ and its perturbed version $\mathbf{A}^{(p)}$, we calculate

$$\alpha_j = \mathbf{u}_j \cdot \mathbf{u}_j^{(p)} \tag{2}$$

$$\beta_j = \mathbf{v}_j \cdot \mathbf{v}_j^{(p)} \tag{3}$$

where α_j ($j = 1$ to z) represents the dot product between the unperturbed and the perturbed jth left ($\mathbf{u}_j$ and $\mathbf{u}_j^{(p)}$) singular vectors and β_j denotes that for the right singular vectors ($\mathbf{v}_j$ and $\mathbf{v}_j^{(p)}$).

Let

$$x_j = \alpha_j + \beta_j \qquad (4)$$

The feature vector for an image is defined as $\mathbf{x} = \{x_j\}$, whose elements quantify the structural distortions in images. The proposed feature vector gives a clear physical meaning of structural degradations in images since changes in $\mathbf{U}$ and $\mathbf{V}$ effectively account for the loss of structural information and hence the perceptual quality as visually exemplified in Fig. 1.

C. Feature Fusion With SVR

As aforementioned, for effective image quality prediction, not only the feature detection is essential but the associated feature fusion procedure also plays an important role. In this work, we formulate image quality prediction as a regression problem based on the proposed features and use SVR to find a mapping function between the features and quality score.

Suppose that $\mathbf{x}_p$ is the feature vector of the pth image in the training image set ($p = 1, 2 \ldots p_m$; p_m is the number of training images). In ε-SVR [14], the goal is to find a function $f(\mathbf{x}_p)$ that has the deviation of ε at most from the actually obtained targets y_p (being the corresponding subjective quality score in this work) for all the training data, and at the same time is as flat as possible. The function to be learned is

$$f(\mathbf{x}) = \mathbf{w}^T \varphi(\mathbf{x}) + \gamma \qquad (5)$$

where $\varphi(\mathbf{x})$ is a nonlinear function of feature vector $\mathbf{x}$, $\mathbf{w}$ is the weight vector, and γ is the bias term. The aim is to find the unknowns $\mathbf{w}$ and γ from the training data such that the error is less than a predefined value denoted by ε.

In the training phase, the SVR system is presented with the training set $\{\mathbf{x}_p, y_p\}$, and the unknowns $\mathbf{w}$ and γ are estimated to obtain the desired function in (5). During the test phase, the trained system is presented with the test vector $\mathbf{x}_q$ of the qth test image and it predicts the estimated objective score y_q ($q = 1$ to q_m; with q_m being the number of test images). We have used the radial basis function (RBF) as the kernel function which is of the form $K(\mathbf{x}_i, \mathbf{x}) = \exp(-\rho \|\mathbf{x}_i - \mathbf{x}\|^2)$ where ρ is a positive parameter controlling the radius. The tradeoff error [14] parameters ρ and ε were determined by using a validation set.

III. SYSTEM PERFORMANCE

A. Databases

For the experiments, we have used three publicly accessible databases in this study: LIVE database [16], IVC database [17], and Toyama database [19]. The LIVE database includes 29 original images from which 779 distorted images were obtained with five types of distortions: fast fading (FF), Gaussian blur, white Gaussian noise (WGN), JPEG, and JPEG2000 (JP2K). The Toyama database contains 182 images out of which 14 are the original images. The rest of the images were JPEG and JP2K coded images (i.e., 84 compressed images for each type of distortion). Six quality scales and six compression ratios were, respectively, selected for the JPEG and JP2K coded images. The following codec softwares were used to generate the compressed images: JPEG using cjpeg software, and JP2K with JasPer software. The IVC database consists of ten original images from which 185 distorted images have been generated, using four different processes: JPEG compression, JP2K compression, locally adaptive resolution (LAR) coding, and blurring. The subjective ratings of the distorted

images have also been provided in all these databases. Thus, a total of 1132 images with varied types and amounts of distortions have been tested as a way to demonstrate the general applicability of the proposed scheme. The distortion types used in these databases often occur in real practical applications.

B. Test Methodology

We have employed the tenfold cross-validation (CV) strategy to test the proposed approach except for the case of testing blurred images from the IVC database as explained later in the text; the data is split into ten chunks; one chunk is used for testing and the remaining nine chunks are used for training. The experiment is repeated with each of the ten chunks used for testing. The average accuracy of the tests over the ten chunks is taken as the performance measure. A nonlinear mapping between the objective model outputs and the subjective quality ratings was employed following the validation method in [18]; we have fitted the objective scores to the subjective scores via a four-parameter cubic polynomial $a_1 x^3 + a_2 x^2 + a_3 x + a_4$ where a_1, a_2, a_3, and a_4 are determined by using the subjective scores and the model outputs.

As shown in Fig. 1, structural degradations are well captured by singular vectors in line with the human perception (for example, in the first row of Fig. 1, the blurred image is of lowest visual quality among the three images and thus its corresponding structure is also more degraded). The effectiveness of feature selection is the basis of the good overall performance of the proposed metric (denoted by Q) in the experimental results that follow.

C. Comparison With Existing Relevant Metrics

We have compared the performance of the proposed system with the most relevant existing image quality estimators, namely, SSIM [4], MSVD [13], and the wavelet-based method VSNR [15]. The experimental results are reported in terms of the three criteria [18] used for performance comparison, namely: Pearson linear correlation coefficient C_P (for prediction accuracy), Spearman rank order correlation coefficient C_S (for monotonicity), and root mean squared error (RMSE), between the subjective score and the objective prediction. For a perfect match between the objective and subjective scores, $C_P = C_S = 1$ and RMSE $= 0$. We can see from the experimental results in Table I that the proposed approach performs better than the existing algorithms by a considerable margin (as indicated by boldfaced values in Table I), in terms of correlation with the subjective scores.

To assess the statistical significance of each metric's performance relative to the other metrics, an F-test was performed on the prediction residuals between the objective predictions (after nonlinear mapping) and the subjective scores. Obviously smaller the residuals, the better the metric is. The test is based on an assumption of Gaussianity of the residual differences. Let σ_A^2 and σ_B^2 denote the variances of the residuals from metrics A and B, respectively, then the F-statistic with respect to metric B is given by $F = \sigma_A^2 / \sigma_B^2$. When $F > F_{\text{critical}}$ where F_{critical} is computed based on the number of residuals and the confidence level, metric A has significantly larger residuals than metric B at a given confidence level. The F_{critical} values for a 99% confidence level are given in Table I. Values of F shown in boldface in Table I signify the cases with 99% confidence for the corresponding metric to have significantly larger residuals than Q. Thus, the proposed metric Q is found to be statistically superior to the existing metrics due to significantly smaller residuals than the other metrics.

We have also computed the C_p values for individual distortion types which are reported in Table II. We can see from Table II that Q loses only in the case of blur distortion testing of the IVC database. The reason for Q losing in this case is most probably due to inadequate training of the model since there are only 20 blurred images in the IVC database (as compared to 174 blurred images in the LIVE database).

TABLE I

C_P, C_S, AND RMSE BETWEEN SUBJECTIVE RATINGS AND TRANSFORMED METRIC OUTPUTS. ALSO LISTED IS THE F STATISTIC FOR EACH METRIC'S RESIDUALS TESTED AGAINST Q'S RESIDUALS. VALUES OF F SHOWN IN BOLDFACE SIGNIFY THAT WITH 99% CONFIDENCE THE METRIC HAS SIGNIFICANTLY LARGER RESIDUALS THAN Q

Measure	Metric	LIVE	Toyama	IVC
C_P	VSNR	0.8894	0.8689	0.7980
	SSIM	0.9117	0.7971	0.7769
	MSVD	0.9094	0.8148	0.7732
	Q	**0.9510**	**0.8997**	**0.8509**
C_S	VSNR	0.8889	0.8599	0.7982
	SSIM	0.9095	0.7871	0.7780
	MSVD	0.8856	0.8146	0.7737
	Q	**0.9365**	**0.8804**	**0.8343**
RMSE	VSNR	8.5674	0.6195	0.7342
	SSIM	9.3365	0.7556	0.7670
	MSVD	11.439	0.7255	0.7726
	Q	**7.0609**	**0.5692**	**0.5988**
F statistic (F)	VSNR	**1.3905**	1.1844	**1.5258**
	SSIM	**1.6513**	**1.7624**	**1.6652**
	MSVD	**2.1978**	**1.6247**	**1.6896**
	Q	1	1	1
$F_{critical}$		1.18	1.43	1.41

TABLE II

PERFORMANCE COMPARISON FOR DIFFERENT DISTORTION TYPES

	Metrics/ Distortions	Q	MSVD	VSNR	SSIM
LIVE	FF	0.954	0.918	0.901	0.938
	Blur	0.951	0.935	0.932	0.946
	JPEG	0.947	0.936	0.944	0.943
	JP2K	0.956	0.924	0.951	0.952
	WGN	0.989	0.930	0.977	0.970
Toyama	JPEG	0.856	0.681	0.801	0.653
	JP2K	0.945	0.911	0.923	0.917
IVC	Blur	0.872	0.824	0.964	0.886
	JPEG	0.822	0.761	0.760	0.740
	JP2K	0.931	0.901	0.861	0.853
	LAR	0.856	0.738	0.737	0.748

Therefore, we have reported a fivefold CV testing result (since for tenfold CV there will be only two images for each test chunk which will be trivial for calculating C_P and C_S) only for this case. Thus, Q generally outperforms the other metrics and achieves better correlation with the subjective scores.

D. Robustness to Untrained Distortions

We have also tested the proposed metric's efficiency in predicting the quality scores of images with a particular type of distortion when the training is done without including any image with that particular distortion. For instance, for the LIVE database, we tested the system in predicting the scores of Fastfaded images given the training is done only with JPEG, JP2K, WGN, and Gaussian blur images. The experiment was repeated for each distortion type. Similar tests were also conducted for IVC and Toyama databases and the C_p values are given in Table III, which indicates that singular vectors are fundamental in quantifying distortions in images since training with the distortion type is not necessary. We can see that C_p values in Table III are quite close to

TABLE III

PERFORMANCE OF Q ON UNTRAINED DISTORTIONS

Test images	LIVE	Toyama	IVC
FF	0.934	-	-
Blur	0.931	-	0.852
JPEG	0.940	0.826	0.787
JP2K	0.946	0.922	0.925
WGN	0.935	-	-
LAR	-	-	0.844

TABLE IV

CROSS-DATABASE PERFORMANCE OF Q

Model/ Test Database	Model$_{LIVE}$	Model$_{IVC}$	Model$_{Toyama}$
LIVE	-	0.901	0.899
IVC	0.841	-	0.802
Toyama	0.878	0.849	-

the C_p values of Q in Table II, i.e., the SVR system predicts the quality effectively without any prior knowledge of the type of distortion.

Q performs better than SSIM, VSNR, and MSVD for majority of the distortions across the three databases even with untrained distortions. Recall that none of the distortion types used in the testing was used in the training. The major reason for the robustness of the proposed system to untrained distortions is that structural degradation is the basic and dominant cause in picture quality degradation which is effectively captured by the changes in **U** and **V**. The advantage of using SVR is that it establishes an optimal and generalized mapping of the features to the perceptual quality score irrespective of the distortion type. This is crucial from a practical point of view due to two reasons: 1) in practice, it is often unfeasible or prohibitive to foresee the type of distortions that caused quality degradation; and 2) quality metrics should be robust regarding the different kinds of distortions in order to benchmark image processing systems [9].

E. Cross-Database Validation

Another way to determine the generality of a machine-learning-based image quality predictor is the cross-database validation since images and/or distortions vary across databases. For the cross-database testing, we used all the images from one database and use the resultant metric to test all the images from the other two remaining databases. For example, we develop a model using the LIVE (denoting it by Model$_{LIVE}$) database and use it to test all the images from the IVC and the Toyama databases. Similarly, we develop Model$_{IVC}$ and Model$_{Toyama}$ and the results for the cross-database testing are presented in Table IV (C_p values are shown). It may be noted that none of the images nor distortions used in testing were used in the training. We can see that the proposed method performs well as indicated by the C_p values. It can be mentioned here that the existing image quality schemes of VSNR, SSIM, and MSVD do not use machine learning. So there is no training procedure involved for these existing schemes and hence cross-database validation does not come into picture for these schemes.

For the IVC database, both Model$_{LIVE}$ and Model$_{Toyama}$ in fact give better prediction accuracies than VSNR, SSIM, and MSVD, even when no IVC database images are used in training. For the LIVE database, Model$_{IVC}$ and Model$_{Toyama}$ perform almost equivalently and give correlation coefficients of 0.901 and 0.899, respectively. For the IVC database, Model$_{LIVE}$ gives correlation coefficient of 0.841 while Model$_{Toyama}$ gives correlation coefficient of 0.802 and so Model$_{LIVE}$

gives higher correlation coefficient compared to $\text{Model}_{\text{Toyama}}$. For the Toyama database, $\text{Model}_{\text{LIVE}}$ gives correlation coefficient of 0.878 while $\text{Model}_{\text{IVC}}$ gives correlation coefficient of 0.849. So $\text{Model}_{\text{LIVE}}$ again gives higher correlation coefficient compared to $\text{Model}_{\text{IVC}}$. Thus, for IVC and Toyama databases, $\text{Model}_{\text{LIVE}}$ has yielded higher correlation coefficient than the other model being compared. This is possibly due to LIVE (with 779 distorted images) being a bigger and more diverse training set with varied image and distortion contents in comparison with the other two databases. Nevertheless, the other two models (i.e., $\text{Model}_{\text{IVC}}$ and $\text{Model}_{\text{Toyama}}$) also perform quite well given their comparatively smaller sizes of training sets. We can see that the C_p values in Table IV compare favorably with those of Q in Table I. These results again confirm that the SVR-based mapping is dependent only on the loss of perceptual quality and not on image and the distortion contents.

F. Further Observations and Discussion

The experimental results demonstrate that the proposed SVR system can learn a generally applicable model and training with close image contents and/or distortion types is not necessary. The databases used in the experiments contain dissimilar images, and the amount and type of distortions are also different (e.g., IVC database contains images with LAR coding defects which LIVE database does not), and so the results prove that the singular vectors are effective features to quantify image quality. These results indicate high generalization ability of the proposed scheme to a large variety of images with varied distortions. The experimental results also show considerable improvement over the relevant existing metrics and this reconfirms the underlying commonality in image quality loss characterized by structural degradations which is the basis of the proposed scheme.

Furthermore, the number of support vectors (SVs) was found to be much smaller than the number of training samples indicating the efficiency of the proposed feature selection and the SVR formulation. The number of SVs is found to decrease with increasing ε value since more samples fall within the ε-tube thereby reducing their number. The experiments show that the number of SVs (for LIVE database) decreases from 555 (for which $C_p = 0.9507$) to as low as 21 (for which $C_p = 0.9446$) with an increase in ε; however, the prediction accuracy does not change appreciably even with just 21 SVs, and this substantiates that quality degradations are well represented by the proposed SVD-based features. This is also significant since ε can be chosen such that the number of SVs is minimal which in turn keeps the computational requirement to a minimum without the loss of prediction accuracy.

For the LIVE database, the linear kernel gave $C_p = 0.86$, polynomial (order 2) kernel yielded $C_p = 0.90$ and as aforementioned the RBF kernel gave $C_p = 0.95$. For the linear kernel, although the C_p is not as high (compared to RBF kernel), it is remarkable because even a simple linear mapping resulted in acceptable performance. The reason for this is that the proposed features can be distinguished reasonably well in the selected feature space.

IV. CONCLUSION

In this brief, we have reported an SVR-based image quality predictor. The major contribution is the adoption and justification of singular vectors as the features to be detected from the images, and the exploitation of SVR for image quality evaluation. A major advantage of using machine learning technique is that the combination of features is optimal since a quantitative data-driven modeling procedure is employed for weight adjustment which results in an appropriate setting of weights for a complex mapping from the feature vector to the desired output. The extensive experimental results confirm that the proposed scheme is more effective, general, and statistically superior in quantifying image quality, than the three relevant existing visual quality metrics. In addition, the proposed SVR-based metric is capable of predicting the perceived quality for images whose content and distortion types do not appear in the training set, as shown in the tests with untrained distortions/databases.

REFERENCES

[1] Z. Wang and A. C. Bovik, *Modern Image Quality Assessment.*. San Rafael, CA: Morgan & Claypool, 2006.

[2] P. Barten, *Contrast Sensitivity of the Human Eye and Its Effects on Image Quality.*. Pittsburgh, PA: SPIE, 1999.

[3] D. M. Rouse and S. S. Hemami, "Analyzing the role of visual structure in the recognition of natural image content with multi-scale SSIM," in *Proc. Western New York Image Process. Workshop*, Oct. 2007.

[4] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error measurement to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[5] W. Lin, L. Dong, and P. Xue, "Visual distortion gauge based on discrimination of noticeable contrast changes," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 15, no. 7, pp. 900–909, Jul. 2005.

[6] A. Blum and P. Langley, "Selection of relevant features and examples in machine learning," *Artif. Intell.*, vol. 97, no. 1-2, pp. 245–271, 1997.

[7] M. P. Eckert and A. P. Bradley, "Perceptual quality metrics applied to still image compression," *Signal Process.*, vol. 70, pp. 177–200, 1998.

[8] P. Gastaldo, S. Rovetta, and R. Zunino, "Objective quality assessment of MPEG-2 video streams by using CBP neural networks," *IEEE Trans. Neural Netw.*, vol. 13, no. 4, pp. 939–947, Jul. 2002.

[9] P. L. Callet, V. G. Christian, and B. Dominique, "A convolutional neural network approach for objective video quality assessment," *IEEE Trans. Neural Netw.*, vol. 17, no. 5, pp. 1316–1327, Sep. 2006.

[10] A. Bouzerdoum, A. Havstad, and A. Beghdadi, "Image quality assessment using a neural network approach," in *Proc. 4th IEEE Int. Symp. Signal Process. Inf. Technol.*, 2004, pp. 330–333.

[11] P. Carrai, I. Heynderickx, P. Gastaldo, R. Zunino, and P. R. Monza, "Image quality assessment by using neural networks," in *Proc. IEEE Int. Symp. Circuits Syst.*, vol. 5, pp. 253–256.

[12] G. W. Stewart, "Stochastic perturbation theory," *SIAM Rev.*, vol. 32, no. 4, pp. 579–610, 1990.

[13] A. M. Eskicioglu, A. Gusev, and A. Shnayderman, "An SVD-based gray-scale image quality measure for local and global assessment," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 422–429, Feb. 2006.

[14] Schölkopf and A. J. Smola, *Learning With Kernels*. Cambridge, MA: MIT Press, 2002.

[15] D. M. Chandler and S. S. Hemami, "VSNR: A wavelet-based visual signal-to-noise ratio for natural images," *IEEE Trans. Image Process.*, vol. 16, no. 9, pp. 2284–2298, Sep. 2007.

[16] H. R. Sheikh, Z. Wang, A. C. Bovik, and L. K. Cormack, "Image and video quality assessment research at LIVE," [Online]. Available: http://live.ece.utexas.edu/research/quality

[17] P. Le Callet, "Florent Autrusseau subjective quality assessment IRCCyN/IVC database," [Online]. Available: http://www2.ir-ccyn.ec-nantes.fr/ivcdb

[18] A. M. Rohaly, J. Libert, P. Corriveau, and A. Webster, Eds., "Final report from the video quality experts group on the validation of objective models of video quality assessment," Mar. 2000 [Online]. Available: www.vqeg.org

[19] Y. Horita, Y. Kawayoke, and Z. M. P. Sazzad, "Image quality evaluation database," [Online]. Available: http://160.26.142.130/toyama_database.zip