IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received November 26, 2018, accepted December 9, 2018, date of publication January 1, 2019, date of current version January 23, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2890304

Blind Stereoscopic Image Quality Assessment Based on Hierarchical Learning

TSUNG-JUNG LIU[1,2], (Member, IEEE), CHING-TI LIN[1,2], HSIN-HUA LIU[3,4], (Student Member, IEEE), AND SOO-CHANG PEI[3,4], (Fellow, IEEE)

[1]Department of Electrical Engineering, National Chung Hsing University, Taichung 40227, Taiwan
[2]Graduate Institute of Communication Engineering, National Chung Hsing University, Taichung 40227, Taiwan
[3]Graduate Institute of Communication Engineering, National Taiwan University, Taipei 10617, Taiwan
[4]Department of Electrical Engineering, National Taiwan University, Taipei 10617, Taiwan

Corresponding author: Hsin-Hua Liu (d02942013@ntu.edu.tw)

This work was supported in part by the Ministry of Science and Technology, Taiwan, under Grant MOST 104-2218-E-005-005 and Grant MOST 105-2221-E-005-062.

ABSTRACT We proposed a blind image quality assessment model which used classification and prediction for three-dimensional (3D) image quality assessment (denoted as CAP-3DIQA) that can automatically evaluate the quality of stereoscopic images. First, in the classification stage, the model separated the distorted images into several subsets according to the types of image distortions. This process will assign the images with the same distortion type to the same group. After the classification stage, the classified distorted image set is fed into the image quality predictor that contains five different perceptual channels which predict the image quality score individually. Finally, we used the regression module of the support vector machine to evaluate the final image quality score, where the input of the regression model is the combination of five channel's outputs. The model, we proposed is tested on three public and popular databases, which are LIVE 3D Image Quality Database Phase I, LIVE 3D Image Quality Database Phase II, and MCL 3D Image Quality Database. The experimental results show that our proposed model leads to significant performance improvement on quality prediction for stereoscopic images compared with other existing state-of-the-art quality metrics.

INDEX TERMS Hierarchical learning, image quality assessment, no reference, stereoscopic images.

I. INTRODUCTION

In recent years, with the development of three-dimensional (3D) technology, more and more multimedia contents of 3D television and 3D movies have been created, and also brought us new experience and challenges. In the process of transmitting, processing and displaying the 3D contents, various distortions affecting image quality perception may be involved. Therefore, it is very necessary to establish an effective method to measure the quality of stereoscopic images. Because Human Visual System (HVS) is the ultimate receiver of images, it is a more direct and accurate way to evaluate the quality of the received images by subjective experience and feeling of human beings. The above assessment method is called subjective image quality assessment (IQA) [1], [2], denoted as SIQA. Since the subjective evaluation of the image quality is directly scored by the human subjects, the perceived quality score closer to the real visual perception of human beings can be obtained. However, this entire assessment process is labor-intensive and time-consuming. During the subjective assessment of image quality, each participant has to do the evaluation in a pre-designed laboratory environment. For example, the viewing distance to the screen, the viewing angle, and the lighting condition of the room has to be kept the same for each participant. Finally, the quality scores of each image obtained from all subjects are averaged to become the final quality score, which is called subjective quality score. There are two ways to calculate subjective quality scores: Mean Opinion Score (MOS) and Differential Mean Opinion Score (DMOS). The MOS is calculated by taking the average of quality scores from each subject for the distorted image, while the DMOS is obtained by the following procedure. First, subtracting the quality score of the corresponding reference image from the quality score of the distorted image given

by each subject to obtain a differential opinion score (DOS). Then the DOSs obtained from all subjects are averaged to produce the DMOS.

From the above description, we can see that the experimental requirements for subjective evaluation of image quality are very troublesome. For instance, a perfect testing environment must be set up first. Second, the subject must focus and concentrate on assessing the quality of the distorted image. The overall evaluation process may take a few days, while the long one can take weeks or even months to establish the subjective quality scores for all the distorted images. Due to the time consumption and inconvenience of subjective IQA, most researchers turn their efforts to objective IQA [1], [2], denoted as OIQA. This type of method builds a metric that can automatically evaluate the quality of distorted images. Over the past decades, many researchers have proposed dozens of methods for objective IQA. One of the classification method for OIQA is based on the availability of reference image information, which can be divided into three types: Full-Reference (FR), Reduced-Reference (RR) and No-Reference (NR) methods, respectively. The FR method requires all information of the reference image as input to assess the quality score of the distorted image. Since this type of method is based on the entire reference image, it generally has better performance. The method of RR is based on partial reference images, so the performance of this method is usually only second to the FR method. The NR method only considers the distorted image as an input. Since there is no reference image as the basis, the performance of this method is usually inferior to the other two. The NR image quality assessment is also referred to as blind IQA.

Due to the presence of depth field information in stereoscopic images, it is not easy to propose an effective objective method to evaluate the quality of distorted stereoscopic images. Up to now, many objective stereoscopic IQA methods have been proposed by researchers and they can be roughly divided into three types.

The first type is the application of two-dimensional (2D) image quality assessment to evaluate the quality scores of stereoscopic images, such as VIF [3], [4], PSNR, SSIM [5], MS-SSIM [6] and so on. Many researchers apply the 2D image quality metrics to the left-view image and the right-view image respectively and then average quality scores from the left and right image to obtain the final stereo image quality score. However, these methods are unsatisfactory in the experimental results of three-dimensional image quality assessment. The second type of stereoscopic image quality assessment takes depth information into account. For instance, Benoit *et al.* [7] proposed a 3D image quality metric which calculated the final stereo image quality score by combining the 2D image quality and the depth image quality. Although there are many studies on the design of image quality assessment for stereoscopic images, the true information of stereoscopic image, such as the disparity image or the cyclopean image, is not yet considered effectively. The third type of stereoscopic image quality assessment is based on the perceptual characteristics of the human visual system. Chen *et al.* [8] proposed a FR 3D IQA method by simulating binocular suppression behavior to solve the problem of binocular rivalry, and to establish an effective model to measure the quality of stereoscopic images. The proposed full-reference stereoscopic IQA method first synthesized the left-view and right-view images into a binocular visual combination image called cyclopean image, and used 2D image quality metrics to evaluate the quality score of the distorted stereo images. The experimental results show that the performance of the proposed structure is superior to the common 2D image quality metrics on asymmetric distorted stereoscopic images. Shao *et al.* [9] proposed a stereo IQA approach based on binocular energy response. These methods have better performances than the previous two types of stereoscopic IQA methods.

Therefore, inspired by previous research [7]–[9], we proposed a blind stereoscopic IQA model which combines the perception information from the monocular images and the binocular visual combination images, and then use the machine learning architecture to evaluate the quality scores of stereoscopic images. Compared with previous works, our proposed method in this paper has the following distinctive features:

1) Using quality score prediction combined with distortion classification help improve the overall performance, especially for cross-database evaluation.
2) In the distortion classification stage, hierarchical binary classifiers are used instead of a multi-class classifier, which is simple and have better classification accuracy than multi-class classification.
3) We use Otsu algorithm to separate the left-view and right-view images into 3 different depth regions, which are the long-range view, the mid-range view, and the close-up view. This can provide us another way to perceive the depth feeling for monocular images.
4) Using different features and regression methods for monocular image (Singular Value Decomposition + Random Forest) and binocular visual combination (BVC) image (Local Binary Pattern + Support Vector Regression) can increase the diversity between these two types of images. In consequence, it increases the possibility to improve the next stage 5-channel (2 monocular image channels plus 3 BVC image channels) fusion performance of predicting visual quality for stereoscopic images.

The rest of the paper is organized as follows. In Section II, we mainly review the related works on stereoscopic image quality assessment and the background of binocular visual combination image. In Section III, we describe the architecture of the proposed 3D image quality assessment. In Section IV, we will describe the experimental results and analyze the performance. Finally, Section V will conclude this paper.

II. RELATED WORKS AND BACKGROUND

In order to explain the proposed objective 3D IQA model in Section III, we give a brief review of previous works and relevant background on binocular visual combination.

A. RELATED WORKS

In this part, we will briefly introduce the stereoscopic IQA methods which are based on machine learning structures. Among these shallow architectures and deep architectures for 3D images, we classify them into two categories, which are: 1) 3D IQA based on a shallow architecture, and 2) 3D IQA using a deep architecture.

In the first category, Chen *et al.* [10] proposed a no-reference stereoscopic IQA approach which extracts 2D and 3D natural scene statistic (NSS) features from stereoscopic images, and then trained by the support vector machine (SVM) regression model to predict the quality score of stereoscopic images. Zhou and Yu [11] proposed a no-reference stereoscopic IQA model that simulates the binocular vision through binocular energy response and binocular rivalry response, and used k-nearest neighbor algorithm to predict the quality score of a stereoscopic image. Shao *et al.* [12] proposed a no-reference stereoscopic IQA structure, which analyzed the correlation between the predictors of the left-view and right-view images, combined the feature-prior and feature-distribution to formulate the stereo image quality prediction, and then used the support vector machine to train the regression model for quality score prediction. Lin *et al.* [13] proposed a no-reference stereo IQA architecture which extracted the commonly used statistical features from stereoscopic images and predicted the stereo image quality scores by training the regression model through SVM.

For the second category, Shao *et al.* [14] proposed a no-reference stereoscopic IQA metric based on the monocular and binocular image interaction. The metric used deep belief network (DBN) to train monocular images and binocular images separately to simulate the human stereo vision in image quality prediction process, and combined both scores to form the final 3D image quality scores through different weighted combinations.

B. BACKGROUND ON HUMAN BINOCULAR VISUAL COMBINATION

Human binocular vision is a very complex visual system and also plays an important role in depth perception. The important clues about the whole system may come from the brightness perception when the human eyes observe the surface of an object. Many researchers have proposed the algorithm that combine two slightly different monocular images into one binocular vision composite image and produce the information of depth perception. Next we will briefly describe the binocular visual composition. First, the binocular visual combination (BVC) image can be synthesized by three different models: eye-weighted model [15], vector summation model [16] and neural network model [17].

In the eye-weighted model, it is based on monocular luminance information or binocular luminance information, forming a linear model for perceiving binocular vision, which is defined as

$$C = [(w_L I_L)^\beta + (w_R I_R)^\beta]^{\frac{1}{\beta}}, \tag{1}$$

where C is the binocular combination image, w_L and w_R are the weighting factors of the left-view and right-view images, I_L and I_R are the left-view and right-view images respectively, and β is the perception parameter, where $\beta = 1$ and 2 represent the monocular luminance and binocular luminance, respectively.

In the vector summation model, the model shows that the perception of binocular luminance is based on the sum of two orthogonal vectors, which is defined as

$$C = [(I_L)^2 + (I_R)^2 + I_L I_R]^{\frac{1}{2}}, \tag{2}$$

where C is the binocular combination image, I_L and I_R are the left-view and right-view images, respectively.

In the neural network model, the model is based on a two-channel model of monocular and binocular images, which is defined as follows

$$C = \frac{I_L}{1 + I_R} + \frac{I_R}{1 + I_L} + 0.1\hat{a}\acute{L}\acute{Z} I_L I_R, \tag{3}$$

where C is the binocular combination image, I_L and I_R are the left-view and right-view images, respectively.

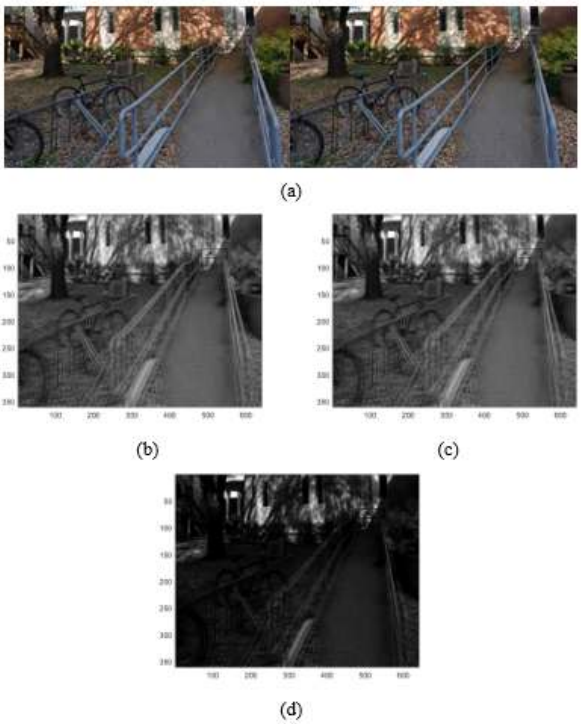


FIGURE 1. Binocular visual combination images. (a) Stereo image pair. (b) The corresponding binocular image based on eye-weighted model. (c) The corresponding binocular image based on vector summation model. (d) The corresponding binocular image based on neural network model.

These three BVC images are shown in Fig. 1 for easy reference.

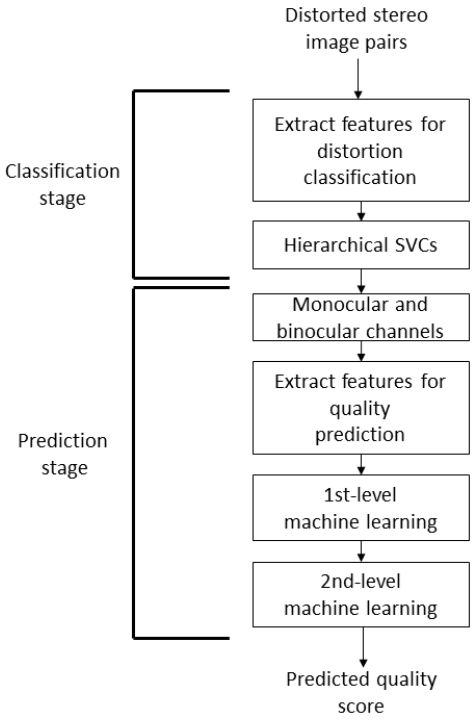


FIGURE 2. The proposed blind stereoscopic image quality assessment model.

III. PROPOSED METHOD

The most effective method to predict the quality of a stereoscopic image is to directly assess the quality of the true binocular visual image which is formed inside the human brain. However, it is difficult to obtain a real binocular visual image. Therefore, in this paper, we propose a NR IQA model that utilizes BVC image and monocular image to imitate human visual system to assess the quality of stereoscopic images. Our proposed model can be roughly divided into two stages, namely, classification stage and prediction stage, as shown in Fig. 2. The details of the blind stereoscopic image quality assessment model that we proposed are described in the following subsections.

A. CLASSIFICATION STAGE

In the classification stage, the overall detailed architecture is shown in Fig. 3. I_d in Fig. 3 represents the distorted stereoscopic image pair, where (I_L, I_R) represents the left-view gray scale image and the right-view gray scale image, respectively. And N_{dis} represents the set of distortion types. The classifiers are in a hierarchical combination. Each classifier only needs to classify the image into 2 types: the image belonged to distortion type n_i, and the image not belonged to distortion type n_i. The reason for us to design these hierarchical classifiers is two-fold: 1) reduce the computational complexity of each classifier, and 2) improve the classification accuracy of each classifier.

1) EXTRACTING FEATURES FOR CLASSIFICATION

We can extract local binary pattern (LBP) [18], [19] features from either gray scale right-view image or left-view image, which is a very useful way to describe the image textures. However, in order to reduce the computation cost in this work, we only extract LBP features from the right-view image instead of both right-view and left-view images. The LBP feature is obtained by using a filter of 3×3 window size for each pixel to compare its grayscale intensity with its neighbor pixels, and then using rotation function [20] to extract the rotational-invariant LBP feature. Here, we set the radius value and number of surrounding neighbors to be equal to 1 and 8 respectively to obtain the rotational-invariant LBP feature, whose total dimension is 36.

2) THE CLASSIFICATION MODEL

The machine learning method we use in the classification stage is support vector classification (SVC), which belongs to one type of SVM. For the implementation, we use LIBSVM package [21] to design a stack of C-SVC classifiers with a radial-basis function (RBF) kernel to classify a given distorted stereoscopic image into one of the distortion types hierarchically [22], [23]. We proposed this hierarchical structure to be able to classify each stereo image into one specific distortion type in order to form several distorted stereo image sets. In the subsequent prediction stage, we design a quality rating system within each type of distortion image set.

B. PREDICTION STAGE

In the prediction stage, we adopt the same quality rating architecture which consists of five different image quality prediction channels for each type of distortion image. The five image quality prediction channels are the left-view image quality prediction channel, the right-view image quality prediction channel, and the three different BVC image quality prediction channels. Fig. 4 shows the block diagram of five image quality prediction channels used in this paper. The left-view and right-view image quality prediction channels can mimic the process of human's monocular vision, while the three different BVC image quality prediction channels can mimic the process of human's binocular vision. Since the quality prediction for left-view and right-view channels are basically the same, except in the input images, we will only describe the evaluation process in right-view image quality prediction channel.

For right-view image quality prediction channel, we firstly transform the right-view image into grayscale image I_R and then divide it into 16 image patches, denoted as $I = \{I_1, I_2, \ldots, I_{16}\}$. Then we use Otsu algorithm [24], [25] to separate each of 16 image patches into 3 different depth regions, which are the long-range view, the mid-range view, and the close-up view. These Otsu filtered image patches are regarded as feature image patches, denoted as $I_{otsu} = [I_{otsu,1}, I_{otsu,2}, \ldots, I_{otsu,16}]$. Next the I_{otsu} is passed through the Sobel filter to generate the edge feature image patches and the corresponding threshold values, denoted as $I_{edge} = [I_{edge,1}, I_{edge,2}, \ldots, I_{edge,16}]$ and $th_{edge} = [th_{edge,1}, th_{edge,2}, \ldots, th_{edge,16}]$, respectively. Then we use

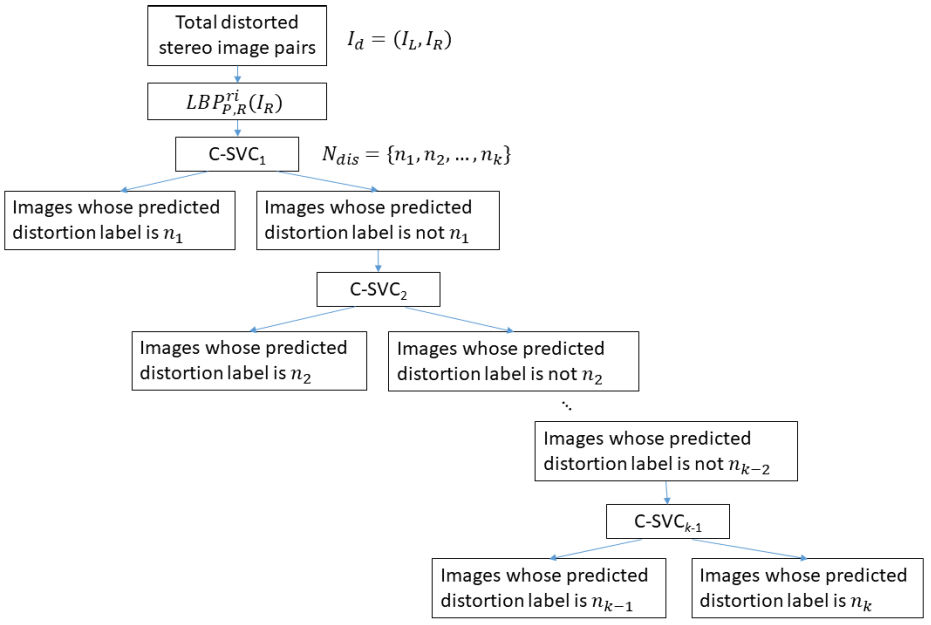


FIGURE 3. Block diagram of the hierarchical classification stage.

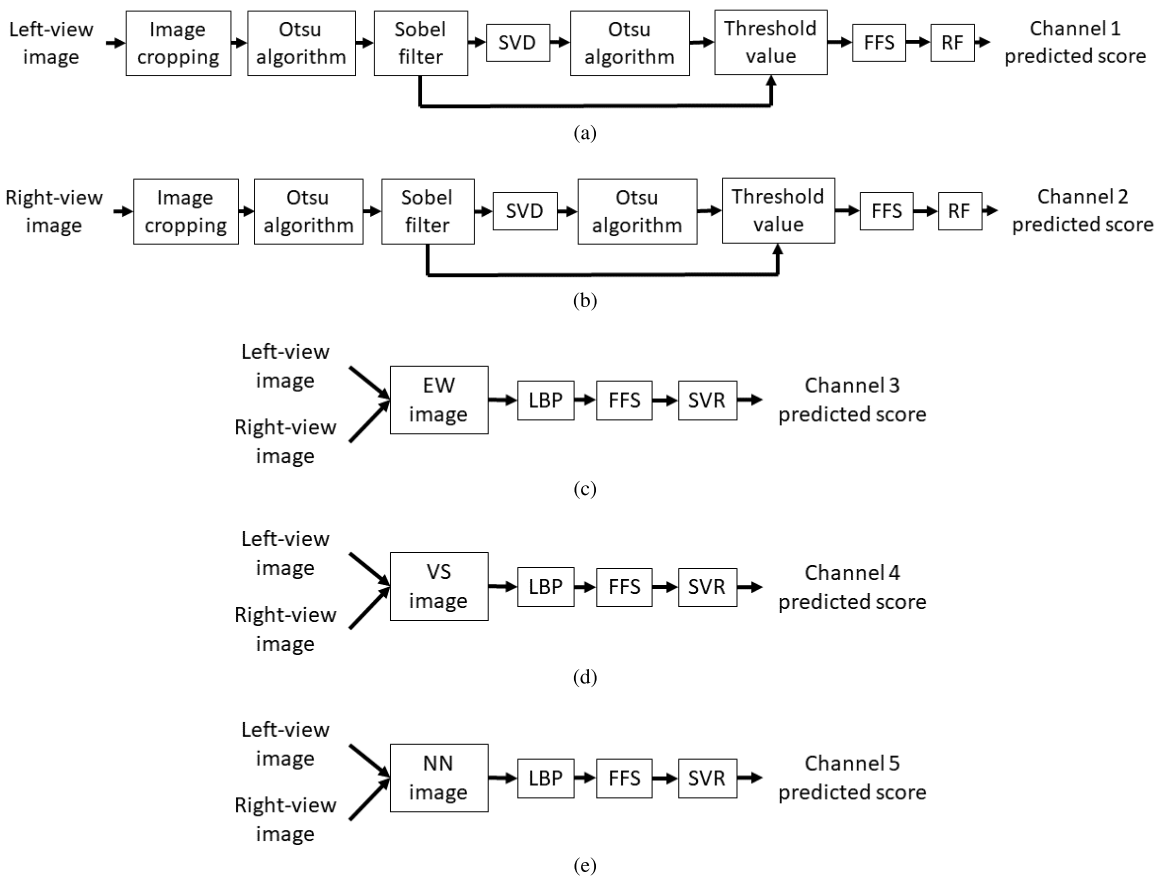


FIGURE 4. The block diagram of five image quality prediction channels used in the prediction stage. (a) Left-view image quality prediction channel, (b) right-view image quality prediction channel, (c) the eye-weighted BVC image quality prediction channel, (d) the vector summation BVC image quality prediction channel, (e) the neural network BVC image quality prediction channel.

singular value decomposition (SVD) [26], [27] to extract the eigenvalues of I_{edge}. After evaluating the eigenvalues of the I_{edge}, we use Otsu algorithm to divide the eigenvalues into 2 groups and obtain the threshold values, denoted as $th_{svd} = [th_{svd,1}, th_{svd,2}, \ldots, th_{svd,16}]$. In Fig. 5, we show the input image and its corresponding image patches created in

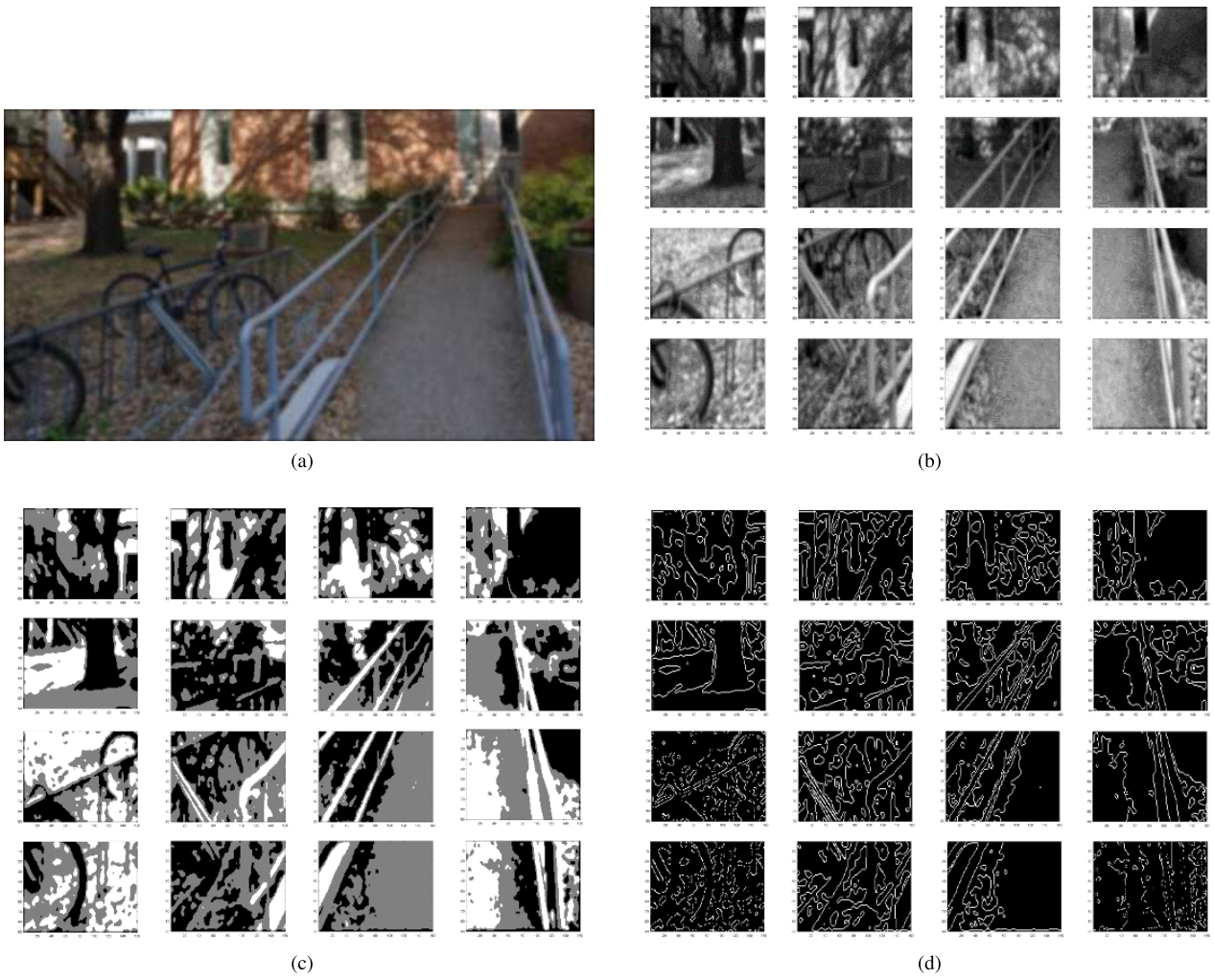


FIGURE 5. Demonstration of image used in the right-view image prediction channel. (a) Right-view image of JPEG distorted stereoscopic image pair, (b) the corresponding image patches I for (a), (c) the respective Otsu filtered image patches I_{otsu}, (d) the respective edge feature image patches I_{edge}.

the intermediate steps of right-view image quality prediction channel.

The two feature vectors of th_{edge} and th_{svd} are combined to form a 32-dimensional right channel feature, denoted as ft_R. Since the size of ft_R is large compared to the size of the dataset, over-fitting is a possibility. Hence, we use forward feature selection (FFS) [28] to reduce the feature dimension. First, we choose the feature that correlates the best with perceptual scores on the training set; then, we find the next feature which combines with the chosen feature in previous step and the combined one has the best correlation with perceptual scores. This process continues until the increment of correlation coefficients is less than 0.001, and the resultant chosen feature combination is denoted as ft'_R. Finally, we use random forest (RF) regression model which consists of 300 decision trees to predict the channel quality score. The predicted quality scores for the left-view and right-view image quality prediction channels are denoted as channel 1 predicted score and channel 2 predicted score, respectively.

In the three BVC image quality prediction channels, the input image is a BVC image that is synthesized from the corresponding left-view and right-view images. According to the image combination model mentioned in Section II, the BVC image can be computed by the following three ways:

1) Eye-Weighted (EW) model

$$C_1 = \left(0.5\left(I_L\left(x,y\right)\right)^2 + 0.5\left(I_R\left(x+d,y\right)\right)^2\right)^{\frac{1}{2}} \quad (4)$$

2) Vector Summation (VS) model

$$C_2 = [\left(I_L\left(x,y\right)\right)^2 + \left(I_R\left(x+d,y\right)\right)^2 + I_L\left(x,y\right)\cdot I_R\left(x+d,y\right)]^{\frac{1}{2}} \quad (5)$$

3) Neural Network (NN) model

$$C_3 = \frac{I_L\left(x,y\right)}{1+I_R\left(x+d,y\right)} + \frac{I_R\left(x+d,y\right)}{1+I_L\left(x,y\right)} + 0.1\cdot I_L\left(x,y\right)\cdot I_R\left(x+d,y\right) \quad (6)$$

Compared with (1)-(3), we know β is set equal to 2, $w_L^2 = w_R^2 = 0.5$, and (x, y) represent the position of an image pixel, d is the disparity index, I_L and I_R represent the left-view image and right-view image, respectively. These three channels are also roughly the same except the input image. So in next paragraph, we will only outline the evaluation procedure in eye-weighted BVC image quality prediction channel.

We first extract LBP features from gray scale eye-weighted BVC image, and the configuration we use for LBP is the same as we did for the classification stage. The extracted feature vector is denoted as ft_{EW}. Then ft_{EW} is fed into FFS process to reduce the feature dimension and becomes ft'_{EW}.

Finally, we use SVM regression model (i.e.,ϵ-SVR with RBF kernel) to acquire the channel predicted score. The predicted quality scores of the EW, VS and NN image quality prediction channels are denoted as channel 3 predicted score, channel 4 predicted score, and channel 5 predicted score, respectively.

To summarize, in this stage, we not only use different features (SVD and LBP) but also use different regression methods (RF and SVR) to learn the prediction model for monocular image and BVC image, respectively. As a result, the diversity between these two types of images will increase and the next stage fusion performance of five quality prediction channels will have larger possibility to be improved.

C. FINAL QUALITY EVALUATION

After obtaining each channel's predicted score, we use the same configuration of SVM regression model [29], [30] to fuse predicted scores from five channels to predict the final stereoscopic image quality score. The procedure is shown in Fig. 6.

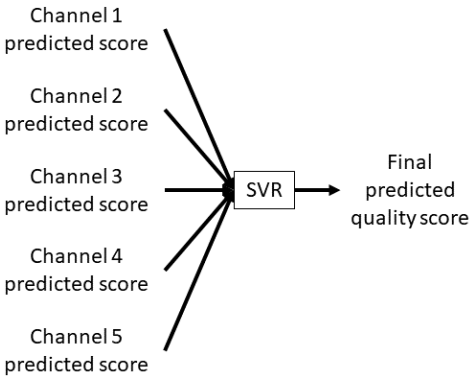


FIGURE 6. The block diagram of final quality evaluation used in the last step of prediction stage.

IV. EXPERIMENTAL RESULTS AND ANALYSIS

The method we presented in the previous Section is called Classification and Prediction 3D Image Quality Assessment (CAP-3DIQA). In this section, we will analyze how well the performance of our proposed image quality metric is in predicting the quality for stereo images. For this purpose, we test the CAP-3DIQA on three public and popular 3D image quality databases, which are LIVE 3D IQA Database Phase I [31], [32], LIVE 3D IQA Database Phase II [33] and MCL 3D IQA Database [34].

A. 3D IQA DATABASES

The LIVE 3D IQA Database Phase I contains 20 reference stereoscopic image pairs and 365 symmetrically distorted stereo image pairs, corresponding to five distortion types: JPEG compression (JPEG) distortion, JPEG2000 compression (JP2K) distortion, additive white noise (WN) distortion, the fast fading (FF) distortion, and the Gaussian blur (Gblur) distortion. While Phase II has the same distortion types as Phase I, eight reference stereoscopic image pairs, and 360 distorted stereo image pairs which include 120 symmetrically and 240 asymmetrically distorted stereo image pairs. Here, if left and right view images both have the same distortion levels, then they will be considered as symmetrically distorted stereo pairs, while they will be considered as asymmetrically distorted stereo pairs if left and right view images have different distortion levels. The subjective quality score is provided in a DMOS value for each distorted stereo image pair in these two 3D IQA databases.

The MCL 3D IQA Database contains 9 reference stereo image pairs and 684 symmetrically distorted stereo image pairs, corresponding to six distortion types, which are JPEG compression (JPEG) distortion, JPEG2000 compression (JP2K) distortion, additive white noise (WN) distortion, Gaussian blur (Gblur) distortion, sampling blur (Sblur), and transmission loss (Tloss). The subjective quality score provided in this 3D IQA database is a MOS value for each distorted stereo image pair.

B. PERFORMANCE MEASURES

In order to benchmark the performance of our proposed model, three popular performance indicators are used. Pearson linear correlation coefficient (PLCC), Spearman rank ordered correlation coefficient (SROCC) and Root Mena Squared Error (RMSE) are used to evaluate the performance of proposed model with respect to the similarity between predicted scores and subjective scores. PLCC and SROCC evaluate the accuracy and monotonicity between the predicted scores and subjective scores, respectively. When PLCC, SROCC are closer to 1 and RMSE is closer to 0, the correlation between the assessment scores and the true scores is better [35].

In all the conducted experiments, each dataset is randomly split into two non-overlapping sets: training set and testing set. After training the model on the training subset, the prediction performance is measured on the test subset. The size of testing set is usually quite small. In each train-test procedure, 80% of the database data is chosen for training, and the remaining 20% is for testing. 1000 iterations of the above training and test procedure are performed by varying the splitting of data over the training and testing sets, and the median results are reported as the final performance.

TABLE 1. Classification accuracy for each hierarchical classifier in three databases.

Database	Classifier	C-SVC$_1$	C-SVC$_2$	C-SVC$_3$	C-SVC$_4$	C-SVC$_5$
LIVE Phase I	Distortion types	JPEG/Not JPEG	JP2K/Not JP2K	WN/Not WN	FF/Gblur	-
	Classification accuracy (%)	100	92.28	100	90.44	-
LIVE Phase II	Distortion types	WN/Not WN	JP2K/Not JP2K	JPEG/Not JPEG	Gblur/FF	-
	Classification accuracy (%)	100	100	100	98.61	-
MCL	Distortion types	WN/Not WN	Gblur/Not Gblur	JP2K/Not JP2K	JPEG/Not JPEG	Sblur/Tloss
	Classification accuracy (%)	91.81	86.64	94.37	86.82	82.38

TABLE 2. Feature dimension results in the prediction stage for each channel in the LIVE 3D IQA Database Phase.

Channel type	Before FFS	After FFS	Reduction percentage of feature dimension (%)
Left-view image	32	19	40.6
Right-view image	32	23	28.1
EW image	36	23	36.1
VS image	36	23	36.1
NN image	36	15	58.3

TABLE 3. Feature dimension results in the prediction stage for each channel in the LIVE 3D IQA Database Phase II.

Channel type	Before FFS	After FFS	Reduction percentage of feature dimension (%)
Left-view image	32	12	62.5
Right-view image	32	12	62.5
EW image	36	29	19.4
VS image	36	27	25.0
NN image	36	24	33.3

In Table 1, we show the classification accuracy for each classifier in three databases. For example, in LIVE Phase I, C-SVC$_1$ first classifies distorted images into two types "JPEG" and "Not JPEG"'. Then C-SVC$_2$ classifies the images belonged to "Not JPEG" into "JP2K" and "Not JP2K". Next, C-SVC$_3$ classifies the images in "Not JP2K" into "WN" and "Not WN". Finally, C-SVC$_4$ classifies images in "Not WN" into "FF" and "Gblur". The classification accuracy rates for C-SVC$_1$ and C-SVC$_3$ are both 100%, while the accuracy rate for C-SVC$_2$ and C-SVC$_4$ are 92.28% and 90.44%. The accuracy rates for all hierarchical classifiers in LIVE Phase II are even higher.

Tables 2-4 list feature dimension results in the prediction stage for each channel (Left-view image, Right-view image, EW image, VS image and NN image) in the LIVE 3D IQA Database Phase I, Phase II and MCL 3D IQA Database, respectively. As we can observe from Table 2, the feature dimension can be reduced by as much as 58.3% for NN image channel and as little as 28.1% for right-view image channel in the LIVE 3D IQA Database Phase I. For LIVE 3D IQA Database Phase II, we can reduce the feature dimension by 62.5% in both left-view and right-view image channels,

TABLE 4. Feature dimension results in the prediction stage for each channel in the MCL 3D IQA Database.

Channel type	Before FFS	After FFS	Reduction percentage of feature dimension (%)
Left-view image	32	21	34.4
Right-view image	32	24	25.0
EW image	36	28	22.2
VS image	36	27	25.0
NN image	36	27	25.0

TABLE 5. Channel prediction performance within CAP-3DIQA for LIVE 3D IQA Database Phase I.

Channel type	PLCC	SROCC	RMSE
Left-view image	0.8797	0.8561	7.8152
Right-view image	0.8883	0.8683	7.5365
EW image	0.9359	0.9304	5.7826
VS image	0.9338	0.9313	5.8731
NN image	**0.9378**	**0.9346**	**5.7051**

TABLE 6. Channel prediction performance within CAP-3DIQA for LIVE 3D IQA Database Phase II.

Channel type	PLCC	SROCC	RMSE
Left-view image	0.7020	0.6677	8,1083
Right-view image	0.6033	0.6090	9.1931
EW image	0.9115	0.9137	4.7344
VS image	0.9051	0.9041	4.9086
NN image	**0.9242**	**0.9273**	**4.3795**

and 19.4% in EW image channel. In MCL 3D IQA Database, the feature dimension has decreased by 34.4% and 22.2% for left-view image channel and EW image channel, respectively. Tables 5-7 show channel prediction performance within CAP-3DIQA for LIVE 3D IQA Database Phase I, LIVE 3D IQA Database Phase II and MCL 3D IQA Database, respectively. The NN image channel has the best prediction performance in LIVE 3D IQA Database Phase I and II, while left-view and EW image channels have better prediction performance than others in MCL 3D IQA Database.

C. OVERALL PERFORMANCE ON INDIVIDUAL DATASET

We first evaluate the FR IQA and NR IQA metrics on each individual dataset. Similar to the proposed CAP-3DIQA, we randomly divide a whole dataset into 80% training subset

TABLE 7. Channel prediction performance within CAP-3DIQA for MCL 3D IQA Database.

Channel type	PLCC	SROCC	RMSE
Left-view image	**0.8738**	0.8772	**1.2790**
Right-view image	0.8615	0.8653	1.3358
EW image	0.8560	**0.8841**	1.3817
VS image	0.8315	0.8700	1.5109
NN image	0.8345	0.8768	1.4914

TABLE 8. Median PLCC, SROCC and RMSE on the LIVE 3D IQA Database Phase I.

Model \ Indicators	PLCC	SROCC	RMSE
Chen [8] (FR)	0.9167	0.9157	6.5503
Chen [10] (NR)	0.8954	0.8912	7.2473
Zhou [11] (NR)	0.9283	0.8875	6.0254
Shao [12] (NR)	0.9531	0.9440	5.1079
Lin [13] (NR)	0.9419	0.9370	5.5190
Shao [14] (NR)	**0.9565**	**0.9449**	**4.7552**
P-3DIQA (NR)	0.9457	0.9397	5.3303
CAP-3DIQA (NR)	0.9412	0.9356	5.5483

TABLE 9. Median PLCC, SROCC and RMSE on the LIVE 3D IQA Database Phase II.

Model \ Indicators	PLCC	SROCC	RMSE
Chen [8] (FR)	0.9065	0.9013	4.7663
Chen [10] (NR)	0.8802	0.8801	5.1024
Zhou [11] (NR)	0.8614	0.8228	5.7591
Shao [12] (NR)	0.9034	0.8849	4.9681
Lin [13] (NR)	0.9130	0.9064	4.6095
Shao [14] (NR)	0.9265	0.9106	4.3381
P-3DIQA (NR)	0.9337	0.9247	4.0567
CAP-3DIQA (NR)	**0.9504**	**0.9492**	**3.5260**

TABLE 10. Median PLCC, SROCC and RMSE on the MCL 3D IQA Database.

Model \ Indicators	PLCC	SROCC	RMSE
Chen [8] (FR)	0.8278	0.8300	1.4596
Shao [14] (NR)	**0.9138**	0.9040	**1.0233**
P-3DIQA (NR)	0.9044	**0.9087**	1.1137
CAP-3DIQA (NR)	0.9001	0.9081	1.1409

TABLE 11. PLCC evaluation on each distortion type for LIVE 3D IQA Database Phase I.

Model	JPEG	JP2K	WN	FF	Gblur
Chen [8]	0.6344	0.9164	**0.9436**	0.7580	0.9417
Chen [10]	0.6952	0.9073	0.9174	0.7349	0.9170
Shao [12]	**0.7843**	0.9275	0.9082	0.8094	0.9032
Lin [13]	0.7613	0.9341	0.9332	0.8610	0.9319
P-3DIQA	0.7459	**0.9420**	0.9395	**0.8855**	**0.9638**
CAP-3DIQA	0.7128	0.9250	0.9298	0.8574	0.9481

and 20% testing subset [36]. Such partition will be performed 1000 times on each dataset to obtain 1000 results on PLCC, SROCC, and RMSE. And the median results are recorded as the final performance measure. The PLCC, SROCC and RMSE results listed in Tables 8, 9 and 10 include the application of proposed CAP-3DIQA, and other well-known metrics on LIVE 3D IQA Database Phase I, Phase II and MCL 3D

TABLE 12. SROCC evaluation on each distortion type for LIVE 3D IQA Database Phase I.

Model	JPEG	JP2K	WN	FF	Gblur
Chen [8]	0.5582	0.8956	**0.9481**	0.6879	**0.9261**
Chen [10]	0.6174	0.8632	0.9189	0.6524	0.8779
Zhou [11]	0.6142	0.8237	0.9146	**0.8674**	0.9162
Shao [12]	0.7596	0.9088	0.9091	0.8186	0.8989
Lin [13]	0.7724	0.9021	0.9293	0.8201	0.9034
Shao [14]	0.6434	0.6253	0.7740	0.8526	0.8927
P-3DIQA	0.7554	**0.9172**	0.9381	0.8403	0.9166
CAP-3DIQA	**0.7616**	0.9008	0.9264	0.8344	0.8912

TABLE 13. PLCC evaluation on each distortion type for LIVE 3D IQA Database Phase II.

Model	JPEG	JP2K	WN	FF	Gblur
Chen [8]	0.8422	0.8426	0.9602	0.9097	0.9650
Chen [10]	**0.9012**	0.8990	0.9473	0.9322	0.9407
Shao [12]	0.8933	0.9013	0.8950	0.9075	0.8904
Lin [13]	0.8124	0.9010	0.8952	0.8957	0.9789
P-3DIQA	0.8908	0.9017	0.9594	**0.9427**	0.9661
CAP-3DIQA	0.8678	**0.9190**	**0.9657**	0.9408	**0.9735**

TABLE 14. SROCC evaluation on each distortion type for LIVE 3D IQA Database Phase II.

Model	JPEG	JP2K	WN	FF	Gblur
Chen [8]	0.8396	0.8344	**0.9554**	0.8890	0.9096
Chen [10]	0.8671	0.8669	0.9520	0.9334	0.9001
Zhou [11]	0.5931	0.7174	0.8910	0.8912	0.9033
Shao [12]	0.8797	0.8833	0.8788	0.8850	0.8817
Lin [13]	0.7740	0.8971	0.9032	0.8806	0.9221
Shao [14]	0.7031	0.9127	0.9658	**0.9387**	0.8722
P-3DIQA	**0.8843**	0.9061	0.9433	0.9258	0.9138
CAP-3DIQA	0.8562	**0.9234**	0.9535	0.9238	**0.9467**

TABLE 15. PLCC evaluation on each distortion type for MCL 3D IQA Database.

Model	JPEG	JP2K	WN	Gblur	Sblur	Tloss
P-3DIQA	**0.9409**	**0.9219**	**0.9135**	0.9479	**0.9530**	0.7618
CAP-3DIQA	0.8779	0.9160	0.8640	**0.9518**	0.9269	**0.7833**

TABLE 16. SROCC evaluation on each distortion type for MCL 3D IQA Database.

Model	JPEG	JP2K	WN	Gblur	Sblur	Tloss
Chen [8]	0.7801	0.8311	0.7630	0.4120	0.8386	0.7606
Shao [14]	0.7992	0.8415	0.6406	0.8993	0.8532	0.5674
P-3DIQA	0.8506	**0.9011**	**0.9256**	**0.9519**	0.9577	**0.7909**
CAP-3DIQA	**0.8512**	0.8908	0.8870	0.9422	**0.9582**	0.7842

IQA Databases, respectively. In addition, we also compare the performance of the proposed method without using the classification stage (denoted as P-3DIQA) in Tables 8 to 10, where the best performance has been highlighted in boldface. Observing Tables 8-10, our proposed methods (P-3DIQA and CAP-3DIQA) are only second to Shao's approach [14] in LIVE 3D IQA Database Phase I compared with other FR and NR models. In MCL 3D IQA Database, P-3DIQA and CAP-3DIQA have better SROCC performance than Shao's method, while they do not perform as well as Shao's method on PLCC and RMSE values. However, in LIVE 3D IQA Database Phase II, the proposed CAP-3DIQA has the best

TABLE 17. Results of cross-database performance.

Trained dataset \ Test dataset	3D IQA model	LIVE-I		LIVE-II		MCL	
		PLCC	SROCC	PLCC	SROCC	PLCC	SROCC
LIVE Phase I	Shao [14]	-	-	0.7791	0.7514	0.5619	0.5403
	P-3DIQA	-	-	0.7872	0.7575	0.4052	0.4124
	CAP-3DIQA	-	-	**0.8669**	**0.8378**	**0.6155**	**0.6136**
LIVE Phase II	Shao [14]	0.8936	0.8917	-	-	0.5889	0.6015
	P-3DIQA	0.8670	0.8562	-	-	0.4619	0.4571
	CAP-3DIQA	**0.9192**	**0.9136**	-	-	**0.6692**	**0.6651**
MCL	Shao [14]	0.6561	0.6347	0.4779	0.4631	-	-
	P-3DIQA	0.2999	0.3381	0.3692	0.3220	-	-
	CAP-3DIQA	**0.7913**	**0.7829**	**0.5546**	**0.5257**	-	-

performance among all competing models, including Shao's. Thus, we can conclude that the proposed approach (CAP-3DIQA) performs the best for asymmetrically distorted stereo image pairs (i.e., LIVE 3D IQA Phase II), and is also competitive with other methods for symmetrically distorted stereo image pairs (i.e., LIVE 3D IQA Phase I and MCL 3D IQA databases). The performance of our model for asymmetrically distorted stereo image pairs can be boosted further by incorporating the classification stage with the prediction stage (i.e., CAP-3DIQA). However, the CAP-3DIQA cannot outperform P-3DIQA in predicting the quality scores of symmetrically distorted stereo image pairs (i.e., LIVE 3D IQA Phase I and MCL 3D IQA databases) since the accuracy rate in the classification stage is not as high as in LIVE 3D IQA Database Phase II.

D. PERFORMANCE ON INDIVIDUAL DISTORTION TYPE

Although in this paper we do not focus on training models for specific distortion type, it is interesting to know the performance on each type of distortion. We test the performance of all IQA models for each distortion type on LIVE 3D IQA Database Phase I, Phase II and MCL 3D IQA Database and list the results in Tables 11 to 16. In LIVE 3D IQA Database Phase I, P-3DIQA ranks the first on PLCC value for JP2K, FF, and Gblur. In LIVE 3D IQA Database Phase II, CAP-3DIQA has the best performance on JP2K, WN, and Gblur. For MCL 3D IQA Database, P-3DIQA performs the best for JP2K, WN, Gblur, and Tloss, while CAP-3DIQA defeats the other 3 models on JPEG and Sblur.

E. CROSS-DATABASE PERFORMANCE EVALUATION

Finally, we perform the cross-database evaluation for the proposed NR 3DIQA models (P-3DIQA, CAP-3DIQA) and Shao's approach [14] (the best 3DIQA method before the appearance of our method). This means we build the trained model on one database, and use the trained model to do the testing on the other two databases. As seen in Table 17, we find the proposed CAP-3DIQA has the best generalization performance in all six scenarios. Although the P-3DIQA does not perform as well as the Shao's method in cross-database evaluation, the improved version (CAP-3DIQA) has overcome this shortcoming by including the classification stage and its performance is better than the metric proposed by Shao [14]. In other words, the CAP-3DIQA outperforms the most state-of-the-art NR 3DIQA approach (Shao's method in [14]) with a significant margin on generalization capability. And the computation time required for our proposed model CAP-3DIQA evaluating a single stereo image pair by running on a computer with Intel CPU Xeon E5-1650 v3 @3.50GHz is around 1.3 seconds, as shown in Table 18, which is fast enough for real-time applications.

TABLE 18. Computation time for CAP-3DIQA.

Model	Elapsed time (secs/image)
CAP-3DIQA	1.3227

V. CONCLUSION

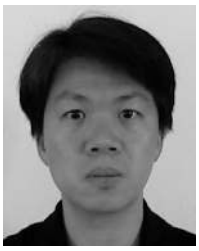
We propose a blind 3D image quality estimator (CAP-3DIQA) to evaluate the quality of distorted stereoscopic images. The model consists of two stages which are classification and prediction. First, the proposed model classifies the distorted image into several subsets according to the distortion types by performing binary classification hierarchically. After the classification stage, the classified distorted stereo image set is passed through the prediction stage, where the prediction of image quality scores is realized via five different prediction channels. Finally, we use a regression model of SVM to fuse the quality scores obtained from five prediction channels to estimate the final image quality. The framework we proposed is tested on three commonly used databases, LIVE 3D Image Quality Database Phase I, Phase II and MCL 3D Image Quality Database, respectively. The experimental results show that the proposed model has significant performance improvement in evaluating the quality of stereoscopic images. Compared with other existing metrics, the proposed model is very competitive in predicting 3D image quality scores among all models and also has the best generalization performance.

REFERENCES

[1] T.-J. Liu, W. Lin, and C.-C. J. Kuo, "Recent developments and future trends in visual quality assessment," in *Proc. Asia–Pacific Signal Inf. Process. Assoc. Annu. Submit Conf.*, 2011, pp. 1–10.

[2] T.-J. Liu, Y.-C. Lin, W. Lin, and C.-C. J. Kuo, "Visual quality assessment: Recent developments, coding applications and future trends," *APSIPA Trans. Signal Inf. Process.*, vol. 2, p. e4, Jan. 2013. [Online]. Available: http://journals.cambridge.org/article_S204877031300005X

[3] H. R. Sheikh and A. C. Bovik, "Image information and visual quality," *IEEE Trans. Image Process.*, vol. 15, no. 2, pp. 430–444, Feb. 2006.
[4] H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A statistical evaluation of recent full reference image quality assessment algorithms," *IEEE Trans. Image Process.*, vol. 15, no. 11, pp. 3440–3451, Nov. 2006.
[5] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.
[6] Z. Wang, E. P. Simoncelli, and A. C. Bovik, "Multiscale structural similarity for image quality assessment," in *Proc. 37th Asilomar Conf. Signals, Syst. Comput.*, vol. 2, Nov. 2003, pp. 1398–1402.
[7] A. Benoit, P. Le Callet, P. Campisi, and R. Cousseau, "Using disparity for quality assessment of stereoscopic images," in *Proc. 15th IEEE Int. Conf. Image Process. (ICIP)*, Oct. 2008, pp. 389–392.
[8] M.-J. Chen, C.-C. Su, D.-K. Kwon, L. K. Cormack, and A. C. Bovik, "Full-reference quality assessment of stereopairs accounting for rivalry," *Signal Process., Image Commun.*, vol. 28, no. 9, pp. 1143–1155, 2013.
[9] F. Shao, G.-Y. Jiang, M. Yu, F. Li, Z. Peng, and R. Fu, "Binocular energy response based quality assessment of stereoscopic images," *Digit. Signal Process.*, vol. 29, no. 1, pp. 45–53, Mar. 2014.
[10] M.-J. Chen, L. K. Cormack, and A. C. Bovik, "No-reference quality assessment of natural stereopairs," *IEEE Trans. Image Process.*, vol. 22, no. 9, pp. 3379–3391, Sep. 2013.
[11] W. Zhou and L. Yu, "Binocular responses for no-reference 3D image quality assessment," *IEEE Trans. Multimedia*, vol. 18, no. 6, pp. 1077–1084, Jun. 2016.
[12] F. Shao, K. Li, W. Lin, G. Jiang, and Q. Dai, "Learning blind quality evaluator for stereoscopic images using joint sparse representation," *IEEE Trans. Multimedia*, vol. 18, no. 10, pp. 2104–2114, Oct. 2016.
[13] C.-T. Lin, T.-J. Liu, and K.-H. Liu, "Visual quality prediction on distorted stereoscopic images," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 3480–3484.
[14] F. Shao, W. Tian, W. Lin, G. Jiang, and Q. Dai, "Toward a blind deep quality evaluator for stereoscopic images based on monocular and binocular interactions," *IEEE Trans. Image Process.*, vol. 25, no. 5, pp. 2059–2074, Mar. 2016.
[15] G. R. Engel, "The visual processes underlying binocular brightness summation," *Vis. Res.*, vol. 7, nos. 9–10, pp. 753–767, 1967.
[16] D. W. Curtis and S. J. Rule, "Binocular processing of brightness information: A vector-sum model," *J. Experim. Psychol., Hum. Perception Perform.*, vol. 4, no. 1, p. 132, 1978.
[17] A. I. Cogan, "Human binocular interaction: Towards a neural model," *Vis. Res.*, vol. 27, no. 12, pp. 2125–2139, 1987.
[18] T. Ojala, M. Pietikäinen, and D. Harwood, "Performance evaluation of texture measures with classification based on Kullback discrimination of distributions," in *Proc. IAPR Int. Conf. Pattern Recognit. (ICPR)*, vol. 1. 1994, pp. 582–585.
[19] T. Ojala and M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on featured distributions," *Pattern Recognit.*, vol. 29, no. 1, pp. 51–59, 1996.
[20] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.
[21] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011.
[22] T.-J. Liu, W. Lin, and C.-C. J. Kuo, "A multi-metric fusion approach to visual quality assessment," in *Proc. 3rd Int. Workshop Qual. Multimedia Exper. (QoMEX)*, 2011, pp. 72–77.
[23] T.-J. Liu, W. Lin, and C.-C. J. Kuo, "Image quality assessment using multi-method fusion," *IEEE Trans. Image Process.*, vol. 22, no. 5, pp. 1793–1807, May 2013.
[24] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.
[25] T.-J. Liu, "Study of visual quality assessment on pattern images: Subjective evaluation and visual saliency effects," *IEEE Access*, vol. 6, no. 1, pp. 61432–61444, 2018.
[26] H. Andrews and C. Patterson, "Singular value decompositions and digital image processing," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-24, no. 1, pp. 26–53, Feb. 1976.
[27] X. Liu, L. Zhang, and K. Lu, "A 3D image quality assessment method based on vector information and SVD of quaternion matrix under cloud computing environment," *IEEE Trans. Cloud Comput.*, to be published.
[28] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, Jan. 2003.
[29] K.-H. Liu, T.-J. Liu, H.-H. Liu, and S.-C. Pei, "Facial makeup detection via selected gradient orientation of entropy information," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2015, pp. 4067–4071.
[30] T. Liu, K. Liu, J. Lin, W. Lin, and C.-C. J. Kuo, "A paraboost method to image quality assessment," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 28, no. 1, pp. 107–121, Jan. 2017.
[31] A. K. Moorthy, C.-C. Su, A. Mittal, and A. C. Bovik, "Subjective evaluation of stereoscopic image quality," *Signal Process., Image Commun.*, vol. 28, no. 8, pp. 870–883, Dec. 2013.
[32] *LIVE 3D Image Quality Database Phase I*. Accessed: Jan. 5, 2019. [Online]. Available: http://live.ece.utexas.edu/research/quality/live_3dimage_phase1.html
[33] *LIVE 3D Image Quality Database Phase II*. Accessed: Jan. 5, 2019. [Online]. Available: http://live.ece.utexas.edu/research/quality/live_3dimage_phase2.htmll
[34] R. Song, H. Ko, and C. C. J. Kuo, "MCL-3D: A database for stereoscopic image quality assessment using 2D-image-plus-depth source," *J. Inf. Sci. Eng.*, vol. 31, no. 5, pp. 1593–1611, 2015.
[35] T.-J. Liu, H.-H. Liu, S.-C. Pei, and K.-H. Liu, "A high-definition diversity-scene database for image quality assessment," *IEEE Access*, vol. 6, no. 1, pp. 45427–45438, 2018.
[36] T.-J. Liu and K.-H. Liu, "No-reference image quality assessment by wide-perceptual-domain scorer ensemble method," *IEEE Trans. Image Process.*, vol. 27, no. 3, pp. 1138–1151, Mar. 2018.



TSUNG-JUNG LIU (S'10–M'14) received the B.S. degree in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 1998, the M.S. degree in communication engineering from National Taiwan University, Taipei, Taiwan, in 2001, and the Ph.D. degree in electrical engineering from the University of Southern California, Los Angeles, CA, USA, in 2014.

He is currently an Assistant Professor with the Graduate Institute of Communication Engineering, and also with the Department of Electrical Engineering, National Chung Hsing University, Taichung, Taiwan. His research interests include computer vision, perceptual image/video processing, visual quality assessment, and big data analytics.

CHING-TI LIN received the B.S. and M.S. degrees in electrical engineering from National Chung Hsing University, Taichung, Taiwan, in 2015 and 2018, respectively.

He is currently in Substitute Military Service with the Taipei Veterans General Hospital, Taitung Branch, Taiwan. His research interests include stereoscopic image quality assessment and machine learning.

HSIN-HUA LIU (S'10) received the B.S. degree in electrical engineering from Feng Chia University, Taichung, Taiwan, in 1996, and the M.S. degree in electrical engineering from National Tsing Hua University, Hsinchu, Taiwan, in 1999.

She is currently pursuing the Ph.D. degree with the Graduate Institute of Communication Engineering, National Taiwan University, Taipei, Taiwan, under the supervision of Prof. S.-C. Pei. Her current research interests include image watermarking, image processing, image retrieval, and color transfer.

SOO-CHANG PEI (S'71–M'86–SM'89–F'00) received the B.S.E.E. degree from National Taiwan University, Taipei, in 1970, and the M.S.E.E. and Ph.D. degrees from the University of California, Santa Barbara, in 1972 and 1975, respectively.

He was an Engineering Officer with the Chinese Navy Shipyard, from 1970 to 1971. From 1971 to 1975, he was a Research Assistant with UCSB. He was a Professor and the Chairman with the Electrical Engineering Department, Tatung Institute of Technology and NTU, from 1981 to 1983, and, from 1995 to 1998, respectively. He is currently a Professor with the Electrical Engineering Department, NTU. His research interests include digital signal processing, image processing, optical information processing, and laser holography.

Dr. Pei is an IEEE Fellow and a member of Eta Keppa Nu and the Optical Society of America. He received the National Sun Yet-Sen Academic Achievement Award in Engineering, in 1984, the Distinguished Research Award from the National Science Council, from 1990 to 1998, the Outstanding Electrical Engineering Professor Award from the Chinese Institute of Electrical Engineering, in 1998, and the Academic Achievement Award in Engineering from the Ministry of Education, in 1998, the IEEE Fellow, in 2000, for contributions to the development of digital Eigen filter design, color image coding, and signal compression, and to electrical engineering education in Taiwan, the Pan Wen-Yuan Distinguished Research Award, in 2002, and the National Chair Professor Award from the Ministry of Education, in 2002 and 2008. He has been the President of the Chinese Image Processing and Pattern Recognition Society, Taiwan, from 1996 to 1998.