4

Reduced-Reference Quality Assessment Methods for Camera and Screen Content Images

In the last two chapters, two different compression methods have been proposed for camera content images (CCIs), and screen content images (SCIs) respectively. In order to analyze the performance of any compression algorithm, it is important to use a reliable image quality assessment (IQA) method. The reliability of an IQA method can be evaluated in terms of its accuracy and robustness with respect to the mean opinion score or subjective score. It was observed from the literature survey in Chapter 1 that although, many IQA methods for CCI's were proposed over the years [Wang and Bovik, 2002; Wang *et al.*, 2004; Chandler and Hemami, 2007; Liu *et al.*, 2012; Zhang *et al.*, 2011] but, due to the distinguishing properties of SCIs compared to the CCIs, such methods did not perform satisfactorily on SCI's. This motivated many research communities to work in the area of IQA for SCIs [Gu *et al.*, 2016b, 2018; Yang *et al.*, 2015; Ni *et al.*, 2016].

Moreover, it was also discussed in Chapter 1 that an RR-IQA method is the most suitable to evaluate the distortion in an image over a communication channel. As the RR-IQA method needs only a few bits of information about the reference image, it can be easily sent to the receiver side in order to verify the quality of the received image. However, it was also observed that the RR-IQA methods perform inferior over distortion caused by compression. Due to the less computation and space complexity in RR-IQA methods, it motivated us to design an efficient IQA metric for CCIs and SCIs. The aim here was to develop an RR-IQA method which can perform efficiently, especially on the distortion caused by compression. This could also justify the performance of the proposed compression framework for SCIs. The major challenge in designing an RR-IQA method was to extract the efficient set of features and descriptors which could correctly represent the reference image at the receiver side. Moreover, a computationally efficient feature matching was also needed in order to provide an accurate IQA metric.

Two feature-based camera and screen-content image quality assessment methods are proposed in this chapter, namely CSQA and FQI. Firstly, the feature points and their descriptors are extracted from reference image (RI) at the sender side and from the distorted image (DI) at the receiver side. A feature matching process is also proposed by evaluating the descriptor distance between every feature of the RI and feature present in the vicinity in the DI. Finally, a normalization process is also proposed in order to find the quality metric which is able to reflect the importance of every feature.

The organization of rest of the chapter is as follows. The proposed CSQA is explained in Section 4.1. The performance analysis of CSQA is carried out in Section 4.2 where its accuracy and robustness is compared with the state-of-the-art methods. Section 4.3 describes the proposed FQI in detail. The result and analysis for FQI are discussed in Section 4.4, where the performance of the proposed FQI is compared with the state-of-the-art methods. Finally, Section 4.5 provides the concluding remarks.

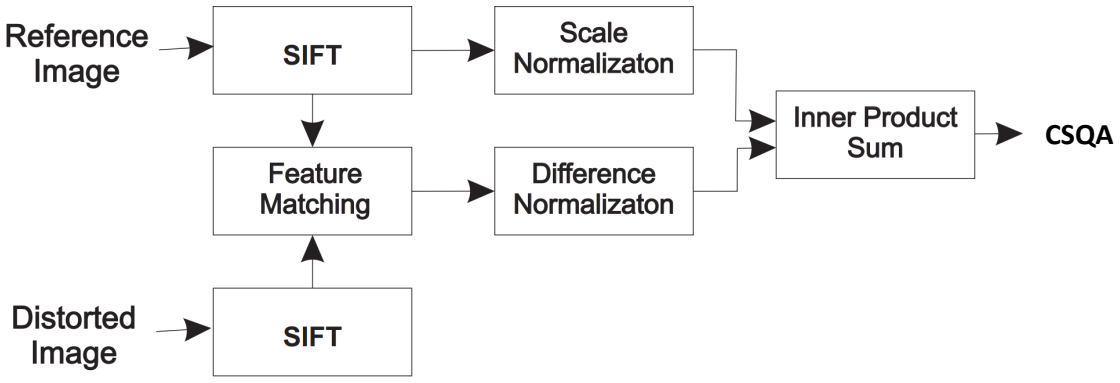


Figure 4.1 : Framework of the proposed Camera and Screen content Image Quality Assessment (CSQA) method.

4.1 CAMERA AND SCREEN CONTENT QUALITY ASSESSMENT: CSQA

The proposed CSQA pipeline is shown in Figure 4.3. SIFT descriptors and the scale values are first extracted for all the feature points in the reference image and the distorted image [Lowe, 2004]. The feature extraction process of SIFT is discussed in detail in the related work section of Chapter 1. To match the features between them, the descriptor distance is calculated between each feature of the reference image and the feature present in the close vicinity in the distorted image. The feature with the least distance is selected as a feature match. The matched feature's distance and their scale values in the reference image reflect the information of the preserved feature in the distorted image and their corresponding importance. To use this as a quality index, these two are normalized as distance vector and scale vector. At last, in order to get the CSQA metric, the inner product between these two vectors is obtained. A detailed description of all the steps are discussed in the following subsections.

4.1.1 Feature Matching

The proposed CSQA, applies SIFT on both the reference and the distorted image in order to get the desired parameters like, location of feature points, the scale values, and the descriptor vectors. To illustrate the feature matching process, let us assume that there are N_1, and N_2 feature points in the reference and distorted image, respectively. Let the location, scale values, and descriptor vector for i^{th} feature point in the reference image are (x_i, y_i), σ_1^i and D_1^i, respectively, where $1 \leq i \leq N_1$. The location, scale values, and descriptor vector for j^{th} feature point in the distorted image are (z_j, w_j), σ_2^j and D_2^j, respectively, where $1 \leq j \leq N_2$. The overall matching process is shown in Algorithm 3, given below. The matched features along with corresponding descriptor distances can be obtained from the vector *Min_Dist* where $Min\_Dist \neq \text{inf}$.

The distortion in SCIs and CCIs are mostly due to noise, blurring, contrast-change, and compression [Yang *et al.*, 2015; Wang *et al.*, 2016; Ponomarenko *et al.*, 2013]. Typically, these distortions don't affect the preserved feature's co-ordinate in the DI with respect to the RI. In view of this fact, to reduce the cost of a typical feature matching process, we propose a fast

and efficient way of feature matching. The proposed feature matching searches for each feature of the RI, preserved in the corresponding close vicinity in the DI. The feature with the lowest Euclidean-distance of the descriptors among the features in the vicinity is selected as the matched features. Then the distances of matched features are saved in vector *Min_Dist*. The vector *Min_Dist* carries two information, such as the features of the RI which are preserved in the DI for all values where *Min_Dist* is not tending to infinity, and the corresponding descriptor distances. Algorithm 3 shows the proposed RDM for feature matching. The length of the vicinity in the DI to be searched for a feature in the RI is set at 2.

Algorithm 3 Proposed Feature Matching Algorithm

```
Min_Dist ← inf
for i = 1 to N1 do
    for j = 1 to N2 do
        if ((|z_j − x_i| ≤ 2) and (|w_j − y_i| ≤ 2)) then
            dist ← Euclidean_Distance(D1^i, D2^j)
            if dist < Min_Dist[i] then
                Min_Dist[i] ← dist
            end if
        end if
    end for
end for
```

So, instead of calculating the Euclidean distance between each feature of the reference image with every feature in the distorted image ($N_1 \times N_2$ distance computations), the proposed feature matching process performs a smaller number of distance computations, as shown in Algorithm 3.

4.1.2 Normalization and Inner-Product

The features in an image have a non-identical importance and the proposed CSQA includes this information to achieve better performance than other feature based quality assessment techniques. As the scale value of a feature point is directly proportional to its gradient-magnitude, we used it to extract the importance of the feature points. In order to reflect the importance of a feature in the proposed CSQA, the scale vector of RI σ_i^1 is normalized and the same is represented as vector S, given in (4.1). As the descriptor distances of the matched features between RI and DI are represented by *Min_Dist* vector, it is normalized in such a way that lower distance carries higher weight-age. The normalization process of descriptor differences is shown in Algorithm 4, where normalized *Min_Dist* vector is represented as T.

$$S[i] = \frac{\sigma_i^1}{\sum_{i=1}^{N_1}(\sigma_i^1)}, \qquad 1 \leq i \leq N_1 \tag{4.1}$$

To reflect the importance of the preserved features in the RI, and their descriptor distances with the DI, the inner product between vector S, and T is computed as shown in (4.14). The inner product in (4.14) provides equal importance to S, and T.

$$CSQA = \langle S, T \rangle = \sum_{i=1}^{N_1}(S[i] \times T[i]) \tag{4.2}$$

Algorithm 4 Descriptor Distance Normalization

for $i = 1$ to N_1 **do**
  $SUM \leftarrow 0$
  if $Min\_Dist[i] \neq \inf$ **then**
    $SUM \leftarrow SUM + Min\_Dist[i]$
      ▷ Variable SUM contains sum of descriptor distances of matched features
  end if
  $T[i] \leftarrow 0$
  if $Min\_Dist[i] \neq \inf$ **then**
    $T[i] \leftarrow 1 - \frac{Min\_Dist[i]}{SUM}$
  end if
end for

4.2 EXPERIMENTAL RESULTS AND DISCUSSIONS

4.2.1 Datasets and Protocol

The performance of the proposed CSQA is evaluated on both CCI and SCI datasets. For CCI, TID-2013 [Ponomarenko *et al.*, 2013], and LIVE [Sheikh *et al.*, 2018] image databases are used. Moreover, we included two publicly available databases for SCI as QACS [Wang *et al.*, 2016], and SIQAD [Yang *et al.*, 2015].

The proposed CSQA is compared with state-of-the-art RR and FR IQA algorithms for both CCIs and SCIs. The RR-IQA methods include WNISM [Wang and Simoncelli, 2005], DNT [Li and Wang, 2009], EPM [Min Zhang, 2011], FTB [Narwaria *et al.*, 2012], SDM [Gu *et al.*, 2013], RIQMC [Gu *et al.*, 2016c] which are proposed specifically CCIs and RRQA [Wang *et al.*, 2018] which has been proposed recently for SCIs. In addition, the FR-IQA algorithms including PSNR [Wang and Bovik, 2002], SSIM [Wang *et al.*, 2004], VSNR [Chandler and Hemami, 2007], GSIM [Liu *et al.*, 2012], VSI [Zhang *et al.*, 2014] which are proposed for CCIs and SQMS [Gu *et al.*, 2016b], SPQA [Yang *et al.*, 2015], GDI [Ni *et al.*, 2016], and SVQI [Gu *et al.*, 2018] which have been specifically proposed for SCIs, are compared as well.

To evaluate the performance of different IQA techniques, four commonly used performance metrics are employed, which includes Spearman rank correlation coefficient (SRCC). Pearson linear correlation coefficient (PLCC), root-mean-squared error (RMSE), and Kendall rank correlation coefficient (KRCC). SRCC and KRCC is used to measure the monotonicity of the prediction of an IQA metric. The major focus of these two performance metrics are on the rank of the data points, and they are invariant of the relative distances between them. On the other hand, PLCC and RMSE is used to evaluate the prediction accuracy by performing a non-linear mapping between the subjective and objective scores as specified in Video-Quality-Experts-Group (VQEG) Phase I FR-TV test [Rohaly *et al.*, 2000]. It should be noted that the correlation value for a few methods on certain dataset was not found and hence left blank.

4.2.2 Performance Evaluation

The validation results on the four SCI and CCI databases in [Yang *et al.*, 2015; Wang *et al.*, 2016; Sheikh *et al.*, 2018; Ponomarenko *et al.*, 2013] are shown in Table 4.1, and Table 4.2. Table 4.1 gives the performance comparison of eight state-of-the-art RR-IQA methods for CCIs, and SCIs with the proposed CSQA. The proposed CSQA outperforms all RR-IQA methods on QACS and LIVE datasets and comes under top 3 for the rest SIQAD, and TID-2013 datasets. CSQA reflected an approximate 6% improvement compared to RWQMS [Wang *et al.*, 2016] which is a recently published IQA method for SCIs. It is observed that the proposed CSQA performs equally well for

Table 4.1 : Performance comparison of CSQA with state-of-the-art RR-IQA methods on CCI and SCI datasets.

	IQA-Model	DNT	EPM	WNISM	FTB	SDM	RIQMC	RRQA	RWQMS	**CSQA Proposed**
SIQAD	PLCC	0.5291	0.6711	0.5857	0.4691	0.6034	0.2732	0.8014	0.8103	0.7928
	SRCC	0.5054	0.6529	0.5188	0.4575	0.6020	0.2395	0.7655	0.7815	0.7695
	RMSE	12.147	10.612	11.602	12.641	11.414	13.769	8.5620	8.8392	8.9218
	KRCC	0.3615	0.4582	0.3540	0.3268	0.4322	0.1627	0.5756	0.5835	0.5924
QACS	PLCC	0.8083	0.6658	0.6326	0.6864	0.6590	0.4241	-	0.8489	0.9162
	SRCC	0.8094	0.6552	0.6154	0.6887	0.7463	0.3489	-	0.8504	0.9201
	RMSE	1.3062	1.6552	1.7182	1.6134	1.6686	2.0091	-	1.1727	0.8902
	KRCC	0.6198	0.4697	0.4352	0.5048	0.5467	0.2502	-	0.6606	0.7723
LIVE	PLCC	0.9173	0.8812	-	0.8968	0.9330	-	-	-	0.9491
	SRCC	0.9287	0.8857	-	0.9073	0.9364	-	-	-	0.9501
	RMSE	-	12.9160	-	12.0863	-	-	-	-	8.5012
	KRCC	-	-	-	0.5611	-	-	-	-	0.8521
TID-2013	PLCC	-	-	-	0.7697	0.5831	0.8651	-	-	0.8523
	SRCC	-	-	-	0.6095	0.3482	0.8044	-	-	0.7895
	RMSE	-	-	-	0.6261	0.7968	0.4920	-	-	0.6309
	KRCC	-	-	-	0.4685	0.2389	0.6178	-	-	0.6480

CCIs and SCIs and shows exceptional stability and consistency.

To further observe the performance of the proposed CSQA, it is also compared with nine state-of-the-are FR-IQA methods for CCIs, and SCIs in Table 4.2. FR-IQA methods are expected to provide better results than RR-IQA because of the availability of RI at the decoder side. It should be noted that the proposed CSQA only requires an average of 0.17 bit-per-pixel as overhead whereas, any FR-IQA method requires 24 bits-per-pixel. The proposed method outperforms all FR-IQA methods on QACS dataset and consistently comes among the top three for LIVE, and TID-2013 datasets. CSQA performed 2.23% better than the second best method SVQI [Gu *et al.*, 2018] which is a recently published work towards FR-IQA for SCIs. This validates the robustness of the proposed CSQA over both CCIs and SCIs. The performance analysis of the proposed CSQA also proves that it performs exceptionally well on QACS dataset which is especially designed with images distorted after compression. The reason for the best performance of CSQA on QACS dataset can be drawn from the fact that while developing an image compression algorithm, it is taken care that the features are preserved in the compressed image to produce better quality reconstructed image.

4.2.3 Dynamic Range Analysis

The dynamic range of an IQA method helps to realize its sensitivity towards change in the distortion level. To analyze the dynamic range of the proposed CSQA, all 29 reference images of the LIVE dataset are compressed using the JPEG baseline and JPEG-2000. For every reference images in the LIVE dataset, 101 JPEG baseline compressed images are obtained with quality factor as 0 (highly compressed), $1, 2, \ldots 100$. We also obtained 99 JPEG 2000 compressed images, between

Table 4.2 : Performance comparison of CSQA with state-of-the-art FR-IQA methods on CCI and SCI datasets.

Dataset	IQA-Model	SSIM	PSNR	VSI	GSIM	VSNR	SQMS	SPQA	GDI	SVQI	**CSQA Proposed**
SIQAD	PLCC	0.5912	0.5869	0.5568	0.5686	0.5966	0.8872	0.8584	0.8515	0.8911	0.7928
	SRCC	0.5836	0.5608	0.5381	0.5483	0.5703	0.8803	0.8416	0.8436	0.8836	0.7695
	RMSE	11.545	11.589	11.890	11.775	11.487	0.6039	7.3421	7.5055	6.4965	8.9218
	KRCC	0.4235	0.4226	0.3874	0.4054	0.4381	0.6936	0.6591	0.6486	0.6985	0.5924
QACS	PLCC	0.8764	0.8669	0.8715	0.8921	0.7050	0.9059	0.8511	0.8669	0.9158	0.9162
	SRCC	0.8829	0.8656	0.8719	0.8947	0.7172	0.9096	0.8456	0.8632	0.9194	0.9201
	RMSE	1.0684	1.1059	1.0879	1.0025	1.5733	0.9396	1.1940	1.1059	0.8909	0.8902
	KRCC	0.7072	0.6768	0.6941	0.7215	0.5383	0.7470	0.6679	0.6812	0.7623	0.7723
LIVE	PLCC	0.9449	0.8723	0.9482	0.9512	0.9213	-	-	-	-	0.9491
	SRCC	0.9479	0.8756	0.9524	0.9561	0.9274	-	-	-	-	0.9501
	RMSE	8.9455	13.3597	8.6816	8.4327	10.5059	-	-	-	-	8.5012
	KRCC	0.7963	0.6865	0.8058	0.8150	0.7616	-	-	-	-	0.8521
TID-2013	PLCC	0.6895	-	0.9000	0.8464	0.7402	-	-	-	-	0.8523
	SRCC	0.6370	0.6395	0.8965	0.7946	0.6812	-	-	-	-	0.7895
	RMSE	0.8608	-	0.5404	0.6603	0.8392	-	-	-	-	0.6309
	KRCC	0.4636	0.4700	0.7183	0.6255	0.5084	-	-	-	-	0.6480

Table 4.3 : Computation cost analysis of proposed CSQA

Feature-Matching Method	Features in RI	Descriptor Dimension	Features in DI	Descriptor Dimension	Distance Calculations	Total Computation
Conventional	$N1$	D	$N2$	D	$N1 \times N2$	$N1 \times N2 \times D$
SIFT	$N1$	128	$N2$	128	$N1 \times N2$	$N1 \times N2 \times 128$
CSQA	$N1$	8	$N2$	8	$N1 \times N2 \times 0.002$	$N1 \times N2 \times 0.256$

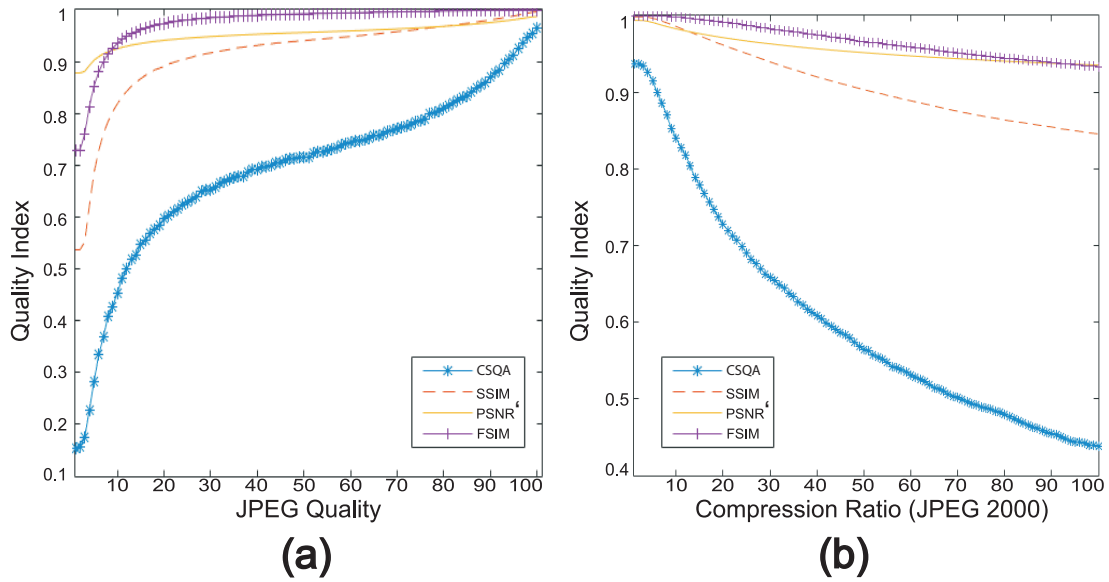


Figure 4.2 : Shows the dynamic range comparison from the plot of quality index of different IQA methods versus JPEG quality in (a), and JPEG 2000 compression-ratio in (b), respectively.

compression-ratio (CR) $2, 3, ..100$ (total $29 \times 200 = 5800$ distorted images).

The proposed CSQA along with PSNR [Wang and Bovik, 2002], SSIM [Wang *et al.*, 2004], and FSIM [Zhang *et al.*, 2011] are applied on these 5800 distorted images. The PSNR index is normalized as $PSNR'$ using (4.3) such that $PSNR' \in [0, 1]$. The results in Figure 4.2 (a) shows the dynamic range after JPEG baseline. The dynamic range for $PSNR'$ comes out to be 0.11 (between 0.88 to 0.99), SSIM 0.45 (0.54 to 0.99), FSIM 0.27 (0.72 to 0.99), and for the proposed CSQA 0.83 (0.15 to 0.98). Similarly, the results in Figure 4.2 (b) shows the dynamic range after JPEG 2000. The dynamic range here for $PSNR'$ is 0.06 (0.93 to 0.99), SSIM is 0.14 (0.85 to 0.99), FSIM is 0.06 (0.92 to 0.99), and for the proposed CSQA is 0.5 (0.43 to 0.93). The observed performance in Figure 4.2 (a), (b) shows significant higher dynamic range of the proposed CSQA as compared to the other methods.

$$PSNR' = 1 - exp^{-(\frac{PSNR}{10})} \tag{4.3}$$

4.2.4 Computation Cost Reduction in Proposed RDM

In order to provide certain invariance in the feature descriptors, the proposed feature extraction process was found to be marginally computationally expensive. The proposed RDM for feature matching tries to recompense the computational cost during feature extraction. The computational cost for RDM is validated on all four datasets for CCI and SCI [Yang *et al.*, 2015; Wang *et al.*, 2016; Sheikh *et al.*, 2018; Ponomarenko *et al.*, 2013]. Typically, feature matching process for N1 features in the RI, and N2 features in the DI requires $(N1 \times N2)$ Euclidean distances to be computed, in order to compare the feature descriptors. The proposed feature matching process of CSQA requires only 0.2% distance computations compared to the conventional feature matching process i.e. $(0.002 \times N1 \times N2)$. Moreover, the key-step in comparing two feature points for a feature matching process is to evaluate the distance of every descriptor between the feature points. The proposed CSQA yields only eight descriptors per feature and hence the computation cost for feature matching is further reduced. Table 4.13 illustrates the cost reduction under proposed CSQA compared to traditional feature matching process where every feature has a D dimensional feature descriptors and SIFT feature matching [Lowe, 2004] where every feature has 128 dimensional descriptors.

4.3 FEATURE QUALITY INDEX: FQI

In the previous section, an RR-IQA method called CSQA was proposed, which showed a promising result on screen content images (SCIs). The performance of CSQA was also found to be satisfactorily on camera content images (CCIs). The aim was to design an IQA method which can conveniently evaluate the quality of the compressed image on the receiver end without requiring large overhead on the communication network. To illustrate such a situation, let us assume that a person is accessing a remote screen over a limited bandwidth network. The reference SCIs are first compressed as per the network bandwidth in order to send the same to the receiving side. In these situations, if a quality assessment can be done on the receiver side in order to analyze the performance of the compression algorithm in real time, the quality of experience (QoE) for the end user can be improved. For CSQA, it was observed that although, its performance was satisfactory on some distortions, the bitrate requirements, and computation costs were on the higher side. Moreover, it was also observed that the performance of CSQA was relatively marginally inferior compared to other state-of-the-art methods on distortions such as contrast-change, and motion-blur.

The reason behind the above-mentioned overhead and computation issue in CSQA was due to the fact that it uses 128 dimensional descriptor for every feature. This high dimensional descriptor requires more computation and overhead for efficient feature matching. The second problem was the inferior performance of CSQA on certain datasets like contrast-change and motion-blur. These distortions were associated with invariance property of SIFT [Lowe, 2004]. We observed that during the weak feature point rejection stage in sift as discussed in Section 1.2.2, if the threshold parameters are made stricter then certain invariance such as contrast-change, and motion-blur could be removed. So, the task to improve the performance of CSQA boiled down to two major challenges. The first challenge was to prune the descriptor size in order to reduce the bitrate requirement in CSQA, and the second challenge was to improve the performance of CSQA on contrast-change, and motion-blur distortions.

With the motivation to solve the above-mentioned issues in CSQA, a superior IQA method called Feature Quality Index (FQI) is proposed through this section. A feature extraction process is proposed in order to extract feature points and their descriptors from RI at the sender side and from the DI at the receiver side. The descriptor size of every feature in FQI is made 16 times smaller than that of CSQA. The feature matching and the normalization process of FQI is on the similar

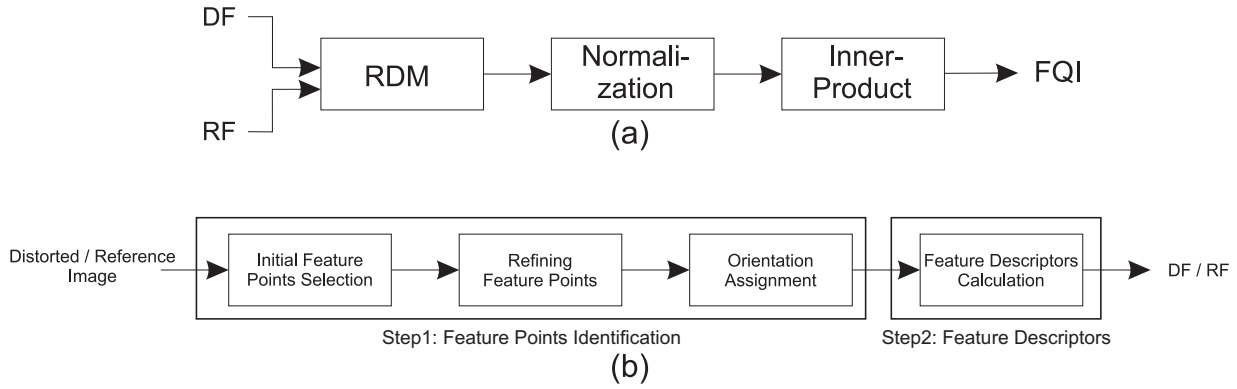


Figure 4.3 : Proposed feature quality index (FQI) framework
(a) FQI framework with proposed feature matching RDM (reduced distance method), (b) Feature extraction process for RI at transmitter Side, DI at Receiver Side

lines of CSQA.

The proposed framework of FQI is shown in Fig. 4.3. In the first stage, shown in Fig. 4.3a, features and descriptors from the reference image (RI) available at the transmitter side, and that of the distorted image (DI) at receiver side are extracted. These will be the reference feature (RF), and distorted feature (DF) respectively. Fig. 4.3b may be referred for feature extraction process. The key steps involved in the feature extraction process are divided into two parts. The first part is responsible for identifying strong feature point and their locations as described in Section 4.3.1, and the second step proposes an eight-dimensional descriptor for every feature points as described in Section 4.3.2. The extracted features RF, and DF are then matched using proposed reduced-distance method (RDM). Finally, FQI is computed as the inner product between the normalized scale values of the matched features and the Euclidean distances. The proposed feature matching and normalization process are discussed in the subsections 4.3.3-4.3.4.

4.3.1 Feature Points Identification

The feature points and their locations in the RI and DI are identified on the lines of Lowe [2004] to include certain invariance to the feature descriptors which is helpful in SCI-IQA. There are three steps involved for this purpose, which are, identification of initial feature points, pruning weak feature points, and orientation assignment. The steps involved in identifying and locating the feature points are described in the following subsections in brief:

Identification of Initial Feature Points

The initial feature points are identified by using cascade filtering approach Wells [1986]. The scale-space or the continuous function of scale, can be useful to identify locations of initial feature points and are also referred to as the Gaussian image. The scale-space of an image $I(x,y)$ is defined as a function $S(x,y,\sigma)$ given in (4.4).

$$S(x,y,\sigma) = G(x,y,\sigma) * I(x,y) \tag{4.4}$$

where $*$ is the convolution operator in x, and y, and

$$G(x,y,\sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \tag{4.5}$$

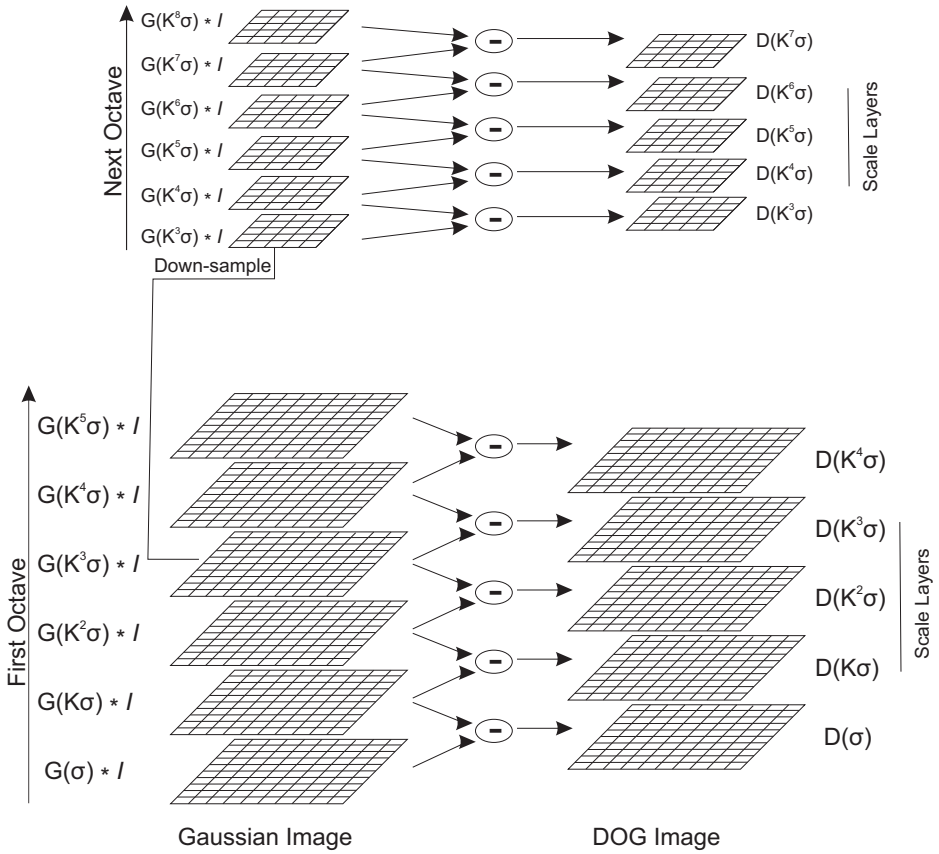


Figure 4.4 : Progressively Gaussian-blurred images (left), and difference-of-Gaussian images (right) are shown for each octave of scale-space. The Gaussian image is down-sampled by a factor of 2 for the next octave.

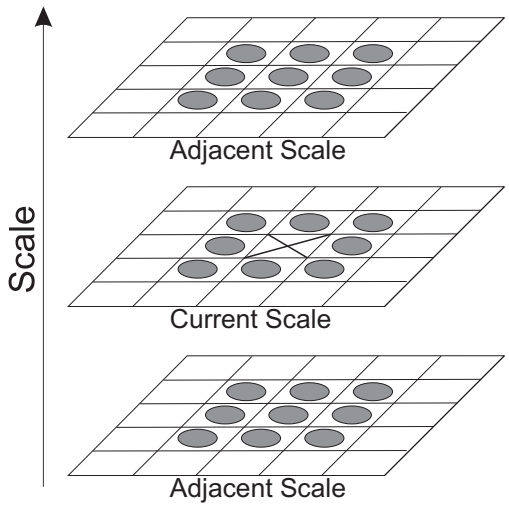


Figure 4.5 : The pixel marked with X is compared to the values at 26 neighbors at the current and adjacent scales (marked with circles) in order to detect the local extrema

To efficiently detect the stable feature points which are invariant of scaling, the difference-of-Gaussian (DOG) images, $D(x, y, \sigma)$ are computed. This is done by evaluating the difference between two nearby scales, which are separated by a constant multiplying factor of k as shown in Fig. 4.4. The DOG images are defined as follows:

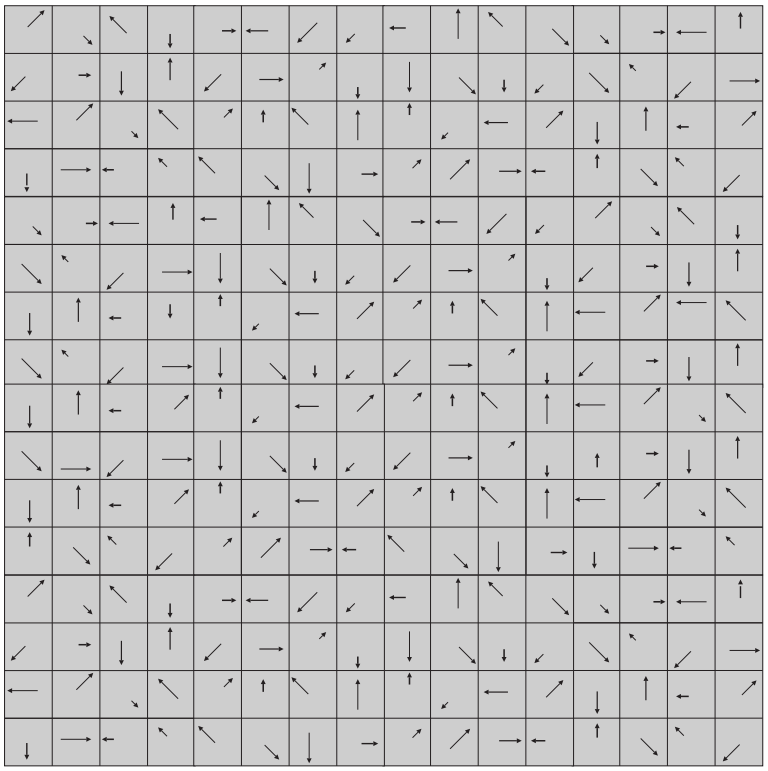
$$\begin{aligned} D(x, y, \sigma) &= (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ &= S(x, y, k\sigma) - S(x, y, \sigma) \end{aligned} \tag{4.6}$$

The scale-space pyramids as shown in Fig. 4.4 are generated for O number of octaves Burt and Adelson [1983]. After each octave, the Gaussian image is down-sampled by a factor of 2, and the process repeated.

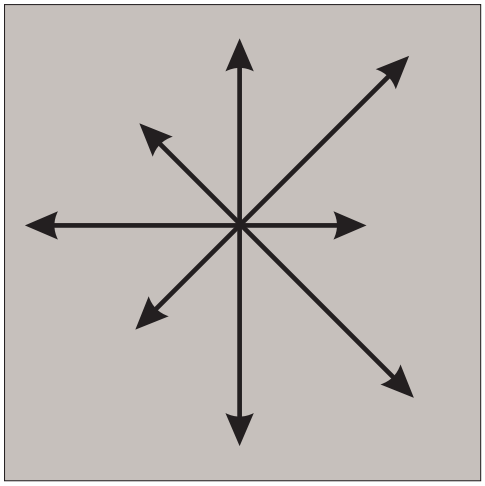
The feature points are then selected by extracting the local extrema from the DOG pyramid as shown in 4.4. For this, Fig. 4.5 can be referred in which pixels at the current scale and those in the two adjacent scales are shown with dark circles and the pixel marked as X, in the current scale, is the current pixel.

Pruning Weak Feature Points

There will be some feature points, obtained from sub-section 4.3.1, that will have low contrast or will be poorly localized along an edge. Such points are referred as unstable feature points. To remove low contrast feature points, the DoG function is first expressed in a small three-dimensional neighborhood around a feature point (x_i, y_i, σ_i) by a second-order Taylor-series as given in 4.7.



(a)



(b)

Figure 4.6 : Evaluating local feature descriptor
(a) Gradient magnitude along with orientation for every feature point in the image around its 16×16 neighborhood after applying a Gaussian filter, (b) The samples in figure (a) are then accumulated into orientation histogram summarizing the content in 1 region with 8 orientations.

$$D(x,y,\sigma) = D(x_i,y_i,\sigma_i) + \left(\frac{\partial D(x_i,y_i,\sigma_i)}{\partial(x,y,\sigma)}\right)^T \Delta + \frac{1}{2}\Delta^T \left(\frac{\partial^2 D(x_i,y_i,\sigma_i)}{\partial(x,y,\sigma)^2}\right)\Delta, \quad \Delta = \begin{pmatrix} x-x_i \\ y-y_i \\ \sigma-\sigma_i \end{pmatrix} \tag{4.7}$$

The location of the extrema, $[\hat{x},\hat{y},\hat{\sigma}]^T$, is then found by taking the derivative of (4.7) with respect to (x_i,y_i,σ_i) and setting it to zero as shown in (4.8). Finally, $D_{extrema}$ is obtained by substituting (4.8) into (4.7) as in (4.9).

$$\begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{\sigma} \end{pmatrix} = -\left(\frac{\partial^2 D(x_i,y_i,\sigma_i)}{\partial(x,y,\sigma)^2}\right)^{-1} \left(\frac{\partial D(x_i,y_i,\sigma_i)}{\partial(x,y,\sigma)}\right) \tag{4.8}$$

$$D_{extrema} = D(x_i,y_i,\sigma_i) + \frac{1}{2}\left(\frac{\partial D(x_i,y_i,\sigma_i)}{\partial(x,y,\sigma)}\right)^T \begin{pmatrix} \hat{x} \\ \hat{y} \\ \hat{\sigma} \end{pmatrix} \tag{4.9}$$

To remove the low contrast feature points, FQI proposed to set the threshold for extrema at 0.06 i.e., if for a feature point $|D_{extrema}| \leq 0.06$, then the same will be discarded. Setting a higher threshold compared to Lowe [2004] yields an average of 3.5% lesser feature points. This helps FQI in achieving lower bitrate in sending the feature information of RI to the receiver. This also helps to improve the performance of FQI under contrast-change distortion as the features become more sensitive towards change in contrast. The remaining feature points give the same results on the quality index as setting the threshold at 0.02 as in Lowe [2004]. However, increasing the threshold value above 0.06 degrades the quality index performance on other distortion categories. The detailed analysis at different threshold value is provided in Section 4.4.2.

Some feature points reside on edges, as edges always give a high response to a DoG filter. These feature points are then eliminated by using Harris corner detector Harris and Stephens [1988]. The principal curvature is computed by using 2×2 Hessian matrix, H at the location and scale of the feature point, as shown in (4.10).

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \tag{4.10}$$

D_{xx}, and D_{yy} are second order derivatives in x, and y respectively. Similarly D_{xy} is derivative in x and then derivative in y, and D_{yx} is derivative in y and then derivative in x. The derivatives are estimated by taking differences of neighboring feature points.

The ratio of eigenvalues of H is proportional to the principal curvature of D. It is observed that a peak which is poorly defined in the DOG function has a large principle curvature along the

edge and small curvature in the perpendicular direction Lowe [2004]. The ratio of eigenvalues is proportional to the ratio of the transpose $Tr(H)$ and the determinant $Det(H)$ of H Harris and Stephens [1988]. The determinant $Det(H)$ as the product of eigenvalues, and transpose $Tr(H)$ as the sum of eigenvalues of H are calculated. The feature point is discarded if the proposed inequality in (4.11) is satisfied Harris and Stephens [1988].

$$\frac{Tr(H)}{Det(H)} \geq 12.5 \tag{4.11}$$

Orientation Assignment

Unlike intensity based IQA techniques such as PSNR or VSNR Wang and Bovik [2002]; Chandler and Hemami [2007], IQA methods are expected to have a rotation invariance Wang *et al.* [2004]. To include this property, FQI Computes the gradient magnitudes and orientations in a small window around the feature point as in Lowe [2004].

For this, an orientation histogram is formed for every feature point by using gradient magnitude $M(x,y)$ and orientation information $\theta(x,y)$ as shown in (4.12), and (4.13), respectively. To do the same, the Gaussian blurred image $L(x,y)$ is produced for every feature points at their corresponding scale value.

$$M(x,y) = \sqrt{[(L(x+1,y) - L(x-1,y))^2 + (L(x,y+1) \overline{-L(x,y-1))^2]}} \tag{4.12}$$

$$\theta(x,y) = \tan^{-1}[(L(x+1,y) - L(x-1,y))/(L(x,y+1) -L(x,y-1))] \tag{4.13}$$

The gradient orientations within the 16×16 neighborhood around the feature point is then used to create an orientation histogram. This histogram has 36 bins corresponding to 360 degree range of orientations. To provide importance to the nearest neighbors of the feature point, a Gaussian-weighted circular window is applied on the gradient magnitude in 16×16 neighborhood with a scaling factor 1.5 times higher than the scale of the feature point. This way, the gradients that are far away from the feature point will add smaller values to the histogram compared to the gradients with the same magnitude which are nearer to the feature points. The samples added to the histogram are weighted by the Gaussian weighted gradient magnitude. The orientation corresponding to the highest peak in the histogram is assigned to the feature point. Also, any peaks above 80% of the highest peak are converted into a new feature point. This new feature point has the same location and scale as the original. But it's orientation is equal to the other peak. So, the orientation assignment can split up one feature point into multiple feature point.

4.3.2 Local Feature Descriptors

The previous subsections assigned the location, scale, and orientation for every feature points in an image. In this subsection, descriptors for every feature points are computed with the aim to make them distinctive as well as invariant of small change in illumination. The feature point descriptors are proposed to be created by computing the gradient magnitude and orientation of sample points around the feature point location with a 16×16 window size. To avoid sudden

changes in the descriptor with small changes in the position of the window, the magnitude of feature points in 16×16 neighborhood are weighted by a Gaussian window. The scaling factor for the Gaussian window is selected as one half of the width of the descriptor window i.e. 8. Choosing this window size also provides less emphasis to gradients that are away from the center of the descriptor, as these are the most affected gradients by registration errors. An orientation histogram is then formed summarizing the contents of 16×16 sized window into 1 region with 8 directions. This leads to a descriptor vector for each feature of length 8.

The Fig. 4.6 can be referred to as an example to get the idea how the descriptors for a feature point are computed. The gradient magnitude along with orientation for every feature point in an image around its 16×16 neighborhood after applying a Gaussian filter, has been shown in Fig. 4.6a. The length and orientation of the arrows are directly proportional to the gradient magnitude and orientation. These are then accumulated into the orientation histogram summarizing the content in 1 region with 8 different directions as shown in Fig. 4.6b. The length of each arrow in Fig. 4.6b corresponds to the sum of the length of arrows in Fig. 4.6a near to that directions.

After getting the descriptors for every feature, each descriptor is quantized to a given number of bits with the aim to fulfil the bit requirements of the network. The analysis of quantizing the descriptor with a different number of bits is given in Section 4.4.2.

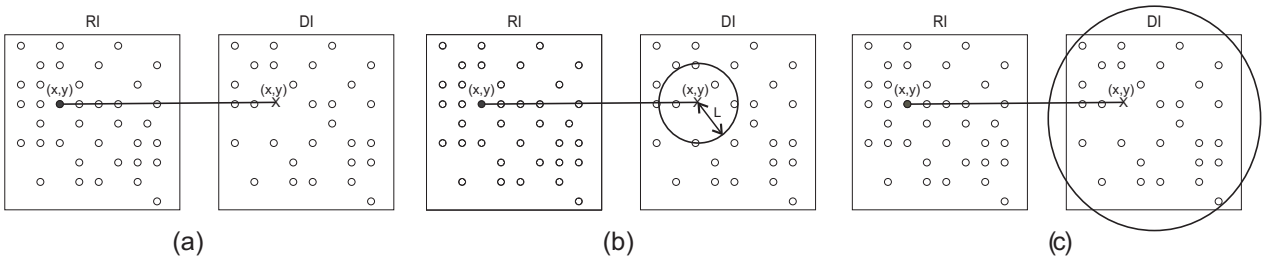


Figure 4.7 : Length of vicinity (L) used in proposed RDM
(a) A feature in Distorted Image (DI) will be searched only at the exact same location in the Reference Image (RI), i.e. FQI_L for $L = 0$, (b) A feature in Distorted Image (DI) will be searched in a length of vicinity L in the Reference Image (RI), i.e. FQI_L, (c) Compare each feature of the RI with every feature of the DI during the feature matching process, i.e. FQI_{all}

4.3.3 Feature Matching: RDM

The feature vector consisting of location, scale, orientation, and descriptors are extracted using the proposed feature extraction method from RI, and DI as RF, and DF respectively. To illustrate the feature matching process, let us assume that there are N_1, and N_2 number of feature points present in RF, and DF, respectively. Let the location, scale, and the descriptor vector for i^{th} feature point in the RF are (x_i^1, y_i^1), σ_i^1, and D_i^1, respectively, where $1 \leq i \leq N_1$. The location, scale, and the descriptor vector for j^{th} feature point in the DF are (x_j^2, y_j^2), σ_j^2, and D_j^2, respectively, where $1 \leq j \leq N_2$.

The distortions in screen content images are usually due to noise, blurring, contrast change, and compression Yang *et al.* [2015]; Wang *et al.* [2016]. Typically, these distortions don't affect the preserved feature's coordinate in the DI with respect to the RI. In view of this fact, to reduce the cost of a typical feature matching process, a fast and efficient way of feature matching is proposed, which will be referred to as reduced-distance method (RDM). The proposed RDM searches for each feature of the RI, preserved in the corresponding close vicinity in the DI. The feature with lowest Euclidean-distance of the descriptors among the features in the vicinity is selected as the matched features. Then, the distances of matched features are saved in vector *Min_Dist*. The vector *Min_Dist*

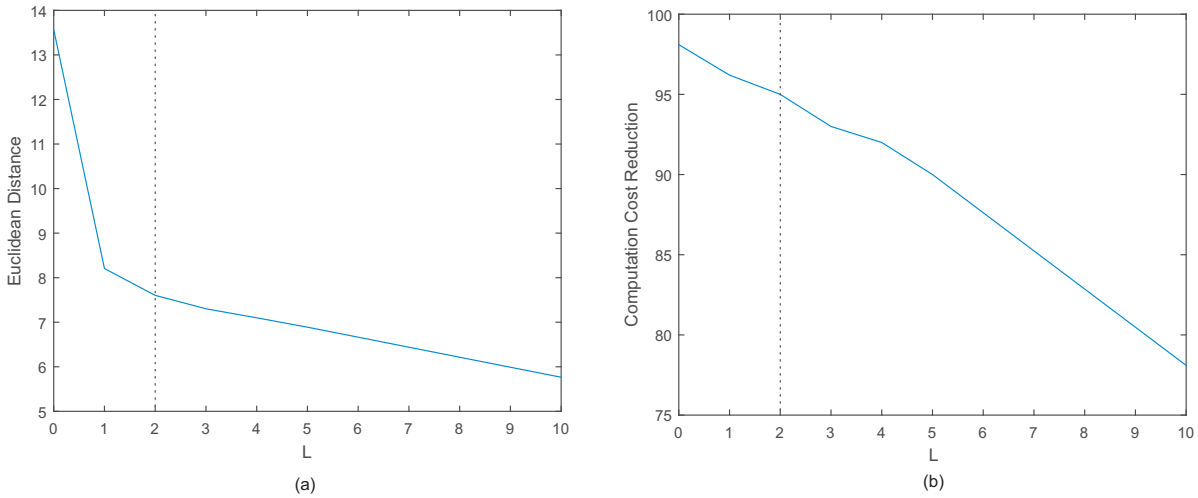


Figure 4.8 : Length of vicinity (L) plot
(a) Euclidean distance plot between every FQI_L and FQI_{all}, (b) Computation cost reduction $(in\%)$ plot for every FQI_L with respect to FQI_{all}

carries two pieces of information, as the features of RI which are preserved in the DI for all values where *Min_Dist* is not tending to infinity, and the corresponding descriptor distances. Algorithm 5 shows the proposed RDM for feature matching. The length of the vicinity in the DI to be searched for a feature in the RI is represented as L.

Algorithm 5 Proposed reduced-distance method (RDM)

```
Min_Dist ← inf
for i = 1 to N1 do
    for j = 1 to N2 do
        if ((|x_j^2 − x_i^1| ≤ L) and (|y_j^2 − y_i^1| ≤ L)) then
            dist ← Euclidean_Distance(D_i^1, D_j^2)
            if dist < Min_Dist[i] then
                Min_Dist[i] ← dist
            end if
        end if
    end for
end for
```

To analyze the optimal length of vicinity (L) to be used in RDM, the proposed FQI was applied on a large-scale screen content images database (SIQAD) Yang *et al.* [2015], and the value of L was varied between $L = [0, 10]$. Here, $L = 0$ means that a feature in DI will be searched only at the exact same location in the RI as given in Fig. 4.7a. For each value of L, the corresponding FQI index (FQI_L) as in Fig. 4.7b, is compared with the FQI values with default feature matching (FQI_{all}) by taking Euclidean distance. Here, the meaning of default feature matching is to compare each feature of the RI with every feature of the DI during the feature matching process as shown in Fig. 4.7c. The curve between $L = [0, 10]$ and the Euclidean distance between FQI_L and FQI_{all} is shown in Fig. 4.8a. The curve between $L = [1, 10]$ and the cost reduction in percentage in terms of the number of comparisons in feature matching is also shown in Fig. 4.8b. The optimal length of vicinity (L) is selected with the aim to get a better trade-off between cost reduction for feature matching, and the Euclidean distance between FQI_L and FQI_{all}. Fig. 4.8a can be approximated as a first degree

Table 4.4 : Benchmark test databases for SCI IQA

Database	Reference Images	Distorted Images	Distortion Types	Image Type	Observers
QACS	24	492	2	Color	20
SIQAD	20	980	7	Color	96

curve having three different slopes between $L = [0,1]$, $L = [1,2]$, and $L = [2,10]$ respectively. The Euclidean distance between $L = [0,2]$ comes down from 13.5 to 7.8 with 2% increase in computation cost, whereas on the other hand, between $L = [2,10]$, the Euclidean distance comes down from 7.8 to 5.9 with 17% increase in computation cost. Setting a threshold on the Euclidean distance between F_L, and F_{all} as 7.8, i.e. $L = 2$ gives a reasonable tradeoff between Euclidean distance and computation cost.

4.3.4 Normalization and Inner-Product

This process is similar to the CSQA normalization process, as discussed in Section 4.1.2. The reason for choosing the same normalization process was the fact that the features in SCIs also possess non-identical importance same as CCIs.

To reflect the importance of the preserved features in the RI, and their descriptor distances with the DI, the inner product between vector S, and T is computed as shown in (4.14). The inner product in (4.14) provides equal importance to S, and T.

$$\begin{aligned} FQI &= \langle S, T \rangle \\ &= \sum_{i=1}^{N_1} (S[i] \times T[i]) \end{aligned} \tag{4.14}$$

4.4 EXPERIMENTAL RESULTS AND DISCUSSIONS

4.4.1 Protocol

To the best of our knowledge, there are two publicly available databases in the SCIs quality assessment (SCI IQA) communities, including QACS Wang *et al.* [2016], and SIQAD Yang *et al.* [2015]. The QACS dataset contains distorted SCI's after HEVC and HEVC-SCC designed for SCI compression and is helpful to analyze the performance of SCI compression techniques. On the other hand, SIQAD contains images distorted after Gaussian noise, Gaussian blur, motion blur, contrast change, and compression. The characteristics of these two databases are provided in Table 4.4. The proposed method is compared with state-of-the-art RR and FR IQA algorithms. The RR-IQA methods include WNISM Wang and Simoncelli [2005], DNT Li and Wang [2009], EPM Min Zhang [2011], FTB Narwaria *et al.* [2012], SDM Gu *et al.* [2013], RIQMC Gu *et al.* [2016c], RWQMS Wang *et al.* [2016], and RRQA Wang *et al.* [2018]. In addition, the FR-IQA algorithms including PSNR Wang and Bovik [2002], SSIM Wang *et al.* [2004], VSNR Chandler and Hemami [2007], GSIM Liu *et al.* [2012], VSI Zhang *et al.* [2014], SQMS Gu *et al.* [2016b], SPQA Yang *et al.* [2015], GDI Ni *et al.* [2016], and SVQI Gu *et al.* [2018] are compared as well. It can be noted that two versions of SSIM implementations are compared, which are denoted as $SSIM_1$ Wang [2018a], and $SSIM_2$ Wang [2018b]. $SSIM_2$ is a single scale version of the $SSIM_1$ where no downsampling has

Table 4.5 : Performance analysis by changing the threshold for $D_{extrema}$ to remove low contrast feature points. The last column represents the average improvement in the performance in terms of percentage.

$D_{extrema}$	SIQAD				QACS				Improvement
Threshold	PLCC	SRCC	RMSE	KRCC	PLCC	SRCC	RMSE	KRCC	(in %)
0.01	0.7086	0.6973	10.1532	0.4863	0.8785	0.8553	1.1265	0.6912	-
0.02	0.7252	0.7198	9.8966	0.5123	0.8873	0.8939	0.9385	0.7294	2.3489
0.03	0.7302	0.7267	9.6456	0.5325	0.8934	0.8983	0.9315	0.7359	1.9585
0.04	0.7403	0.7334	9.6238	0.5417	0.9097	0.9111	0.9214	0.7431	1.8244
0.05	0.7426	0.7386	9.6231	0.5439	0.9102	0.9120	0.9208	0.7441	1.5878
0.06	**0.7471**	**0.7401**	**9.6211**	**0.5481**	**0.9108**	**0.9124**	**0.9201**	**0.7446**	**0.6059**
0.07	0.7470	0.7398	9.9218	0.5478	0.9101	0.9122	0.9211	0.7442	-0.0405
0.08	0.7362	0.7253	10.0112	0.5254	0.8998	0.8976	0.9323	0.7327	-1.4457

Table 4.6 : Performance analysis by changing the number of bits needed to represent a descriptor as bits-per-feature. The first column represents the total number of bits used to send the descriptors of one feature. The overhead in the second column is given in bits-per-pixel (bpp). The last column represents the average improvement in the performance in terms of percentage. The value of $D_{extrema}$ is set at 0.06.

bits-per-feature	Overhead (in bpp)	SIQAD				QACS				Improvement (in %)
		PLCC	SRCC	RMSE	KRCC	PLCC	SRCC	RMSE	KRCC	
48	0.0362	0.7051	0.6932	10.7623	0.4997	0.8926	0.8951	1.3632	0.6993	-
64	0.0968	0.7471	0.7401	9.6211	0.5481	0.9108	0.9124	0.9201	0.7446	6.4779
80	**0.1774**	**0.8018**	**0.7729**	**8.7312**	**0.5803**	**0.9176**	**0.9212**	**0.8897**	**0.7756**	**7.3216**
96	0.2580	0.8024	0.7735	8.5309	0.5838	0.9192	0.9219	0.8838	0.7765	0.6031
112	0.3386	0.8045	0.7743	8.3429	0.5853	0.9208	0.9234	0.8796	0.7798	0.2617
128	0.4192	0.8076	0.7757	8.3395	0.5892	0.9213	0.9249	0.8731	0.7813	0.2611

Table 4.7 : Performance comparison between FQI and different RR-IQA methods on the QACS dataset. The top two performances are highlighted.

IQA-Model	DNT	EPM	WNISM	FTB	SDM	RIQMC	RWQMS	RRQA	CSQA	**FQI Proposed**
PLCC	0.8083	0.6658	0.6326	0.6864	0.6590	0.4241	0.8489	-	**0.9162**	**0.9176**
SRCC	0.8094	0.6552	0.6154	0.6887	0.7463	0.3489	0.8504	-	**0.9201**	**0.9212**
RMSE	1.3062	1.6552	1.7182	1.6134	1.6686	2.0091	1.1727	-	**0.8902**	**0.8897**
KRCC	0.6198	0.4697	0.4352	0.5048	0.5467	0.2502	0.6606	-	**0.7723**	**0.7756**

Table 4.8 : Performance comparison between FQI and different FR-IQA methods on the QACS dataset. The top two performances are highlighted.

IQA-Model	$SSIM_1$	$SSIM_2$	PSNR	VSI	GSIM	VSNR	SQMS	SPQA	GDI	SVQI	CSQA	**FQI Proposed**
PLCC	0.8764	0.8696	0.8669	0.8715	0.8921	0.7050	0.9059	0.8511	0.8669	0.9158	**0.9162**	**0.9176**
SRCC	0.8829	0.8683	0.8656	0.8719	0.8947	0.7172	0.9096	0.8456	0.8632	0.9194	**0.9201**	**0.9212**
RMSE	1.0684	1.0953	1.1059	1.0879	1.0025	1.5733	0.9396	1.1940	1.1059	0.8909	**0.8902**	**0.8897**
KRCC	0.7072	0.6910	0.6768	0.6941	0.7215	0.5383	0.7470	0.6679	0.6812	0.7623	**0.7723**	**0.7756**

Table 4.9 : Performance comparison between FQI and different RR-IQA methods on the SIQAD dataset. The top two performances are highlighted.

IQA-Model	DNT	EPM	WNISM	FTB	SDM	RIQMC	RWQMS	RRQA	CSQA	**FQI Proposed**
PLCC	0.5291	0.6711	0.5857	0.4691	0.6034	0.2732	**0.8103**	0.8014	0.7928	**0.8018**
SRCC	0.5054	0.6529	0.5188	0.4575	0.6020	0.2395	**0.7815**	0.7655	0.7695	**0.7729**
RMSE	12.147	10.612	11.602	12.641	11.414	13.769	8.8392	**8.5620**	8.9218	**8.7312**
KRCC	0.3615	0.4582	0.3540	0.3268	0.4322	0.1627	**0.5835**	0.5756	0.5924	**0.5983**

been performed. $SSIM_2$ is found to be more effective if used at the appropriate scale to preprocess the reference and distorted SCIs.

To evaluate the performance of different IQA techniques, four commonly used performance metrics are employed, including Spearman rank correlation coefficient (SRCC),

Table 4.10 : Performance comparison between FQI and different FR-IQA methods on the SIQAD dataset. The top two performances are highlighted.

IQA-Model	$SSIM_1$	$SSIM_2$	PSNR	VSI	GSIM	VSNR	SQMS	SPQA	GDI	SVQI	CSQA	**FQI Proposed**
PLCC	0.5912	0.7561	0.5869	0.5568	0.5686	0.5966	**0.8872**	0.8584	0.8515	**0.8911**	0.7928	0.8018
SRCC	0.5836	0.7566	0.5608	0.5381	0.5483	0.5703	**0.8803**	0.8416	0.8436	**0.8836**	0.7695	0.7729
RMSE	11.545	9.3676	11.589	11.890	11.775	11.487	**6.6039**	7.3421	7.5055	**6.4965**	8.9218	8.7312
KRCC	0.4235	0.5583	0.4226	0.3874	0.4054	0.4381	**0.6936**	0.6591	0.6486	**0.6985**	0.5924	0.5983

Pearson linear correlation coefficient (PLCC), root-mean-squared error (RMSE), and Kendall rank correlation coefficient (KRCC). SRCC and KRCC can measure the prediction monotonicity of an IQA metric. These two metrics operate only on the rank of the data points and ignore the relative distance between them. On the other hand, PLCC and RMSE evaluate the prediction accuracy by performing a nonlinear mapping between the subjective and objective scores as specified in Video Quality Experts Group (VQEG) Phase I FR-TV test Rohaly *et al.* [2000]. To compute these two metrics, it is needed to apply a regression analysis, as in (4.15). where $\beta_i = 1,2,...,5$ are the parameters to be fitted. A better objective IQA index is expected to achieve higher values in PLCC, SRCC, and KRCC, and lower values in RMSE.

$$f(x) = \beta_1 \left(\frac{1}{2} - \frac{1}{1 + e^{\beta_2 (x - \beta_3)}} \right) + \beta_4 x + \beta_5 \tag{4.15}$$

4.4.2 Parameter Analysis

Threshold Value for Contrast Change

The extrema value ($D_{extrema}$) is used as a threshold to remove any low contrast feature point as given in (4.9). This means, if, for a feature point $|D_{extrema}| \leq Threshold$, then the same will be discarded. A performance analysis is carried out to optimize the threshold value as given in Table 4.5. The performance is observed by evaluating PLCC, SRCC, RMSE, and KRCC on SIQAD and QACS datasets at the given threshold. The last column of the table represents the average improvement in the performance in terms of percentage compared to the performance at the threshold value in the previous row. It has been observed that the performance is optimum with the threshold value equal to 0.06 which validates the statement in Section 4.3.1.

Descriptor Quantization Value

To reduce the transmission overhead and optimize the performance of FQI, the analysis for the number of bits per feature is carried out as discussed in Section 4.3.2. Table 4.6 shows the performance analysis of the proposed FQI by changing the number of bits required to quantize a descriptor and to represent a feature. Every feature has eight descriptors, thus, the total number of bits needed to send a feature will be eight times the number of bits assigned for a descriptor as given in the first column of the table. The performance is observed by evaluating PLCC, SRCC, RMSE, and KRCC on SIQAD and QACS datasets for the given bit-rate. It is observed that the overhead for sending the feature information along with the distorted image, increases linearly, as provided in the second column of the table in terms of bits-per-pixel. The last column in the table

represents the average improvement in the performance in terms of percentage compared to the performance at the number of bits per feature in the previous row. The analysis shows that, by assigning more than 80 bits-per-feature, gives an insignificant improvement in the performance to justify overhead. Therefore, 80 bits-per-feature is found to be appropriate.

4.4.3 Performance Evaluation

The validation results of the two screen content databases in Yang *et al.* [2015]; Wang *et al.* [2016] are shown in Tables 4.7- 4.10. Table 4.7 outlines the performance comparison of proposed FQI and seven state-of-the-art RR-IQA methods for SCI. It is observed that the proposed FQI outperforms all other RR-IQA methods, especially RWQMS Wang *et al.* [2016], which is a recently published IQA method, specially designed for quality assessment for compressed SCI. The proposed FQI reflects an approximate 6.5% improvement compared to RWQMS. This motivated us to compare the performance of the proposed FQI with six FR-IQA methods which are shown in Table 4.8. Due to the availability of RI at the receiver end, it is obvious that FR-IQA methods will yield better results compared to RR-IQA methods where the receiver side only has the feature information of the RI. With this broad understanding of the superiority of FR-IQA, performances were compared to show that the proposed method provides at par performance compared to FR-IQA. The proposed FQI outperforms all FR-IQA methods for QACS dataset Wang *et al.* [2016], especially SQMS Gu *et al.* [2016b], GDI Ni *et al.* [2016], and SVQI Gu *et al.* [2018], which are specifically developed IQA methods for SCIs. The proposed method gives 0.7% better results compared to SVQI Gu *et al.* [2018] on all the images of QACS dataset as shown in Table 4.8.

The reason for which the proposed FQI outperforms all RR and FR-IQA methods can be drawn from the fact that the distorted images in the QACS dataset have undergone compression and while developing an image compression algorithm, it is taken care that the features are preserved in the compressed image to produce better quality reconstructed image Rohaly *et al.* [2000].

The performance of the proposed FQI on SIQAD dataset Yang *et al.* [2015] is shown in Table 4.9 with respect to RR-IQA methods, and 4.10 with respect to FR-IQA methods. It is observed that the proposed FQI perform at second best among the RR-IQA methods in Table 4.9 and it performs at par among FR-IQA methods in Table 4.10. The overall performance analysis also shows the robustness of the proposed FQI as it performs at the best for the QACS dataset and the second best for SIQAD dataset. To evaluate the performance of the proposed FQI further, the analysis of individual distortion type is carried out which is explained in detail in Section 4.4.4.

4.4.4 Performance on Individual Distortion Type

To further evaluate the breakdown prediction performance for individual distortion type, Tables 4.11 and 4.12 can be referred, where the performance of RR-IQA methods are shown for individual distortion present in SIQAD Yang *et al.* [2015], and QACS Wang *et al.* [2016] datasets, respectively. To reflect the prediction accuracy and monotonicity, PLCC and SRCC are used. Similar conclusions could be drawn by using other measures, such as RMSE and KRCC.

Table 4.11 shows the distortion wise performance comparison between FQI and other RR-IQA methods on SIQAD dataset Yang *et al.* [2015]. It should be noted that the compression distortion includes JPEG, JPEG 2000, and Layered compression. The proposed FQI outperforms all other RR-IQA methods under four distortion categories such as Gaussian noise, and three different types of compression. However, for contrast-change, and motion blur distortions the proposed method is second and third to the best method, respectively. Moreover, the proposed FQI performs very well under compression and Gaussian noise categories compared to that of RWQMS Wang *et al.* [2016], and RRQA Wang *et al.* [2018]. The proposed FQI exhibits approximately 2.3% better

Table 4.11 : Performance comparison of FQI with respect to individual distortion type for SIQAD dataset. The top two performances are highlighted.

IQA Model	Gausian Noise		Gausian Blur		Motion Blur		Contrast Change		Compression	
	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC
DNT	0.8189	0.8211	**0.8946**	0.8875	0.7928	0.7903	**0.7846**	**0.6719**	0.5987	0.5867
EPM	0.8420	0.8151	0.8318	0.8293	0.7143	0.7112	0.7151	0.5090	0.5468	0.5168
WNISM	0.8570	0.8442	0.8524	0.8370	0.6618	0.6606	0.7402	0.6142	0.3170	0.2991
FTB	0.7185	0.7165	0.7358	0.7400	0.5984	0.5866	0.5207	0.1112	0.5798	0.5649
SDM	0.8694	0.8635	0.7836	0.8199	0.5434	0.5307	0.7831	0.6617	0.6837	0.6777
RIQMC	0.4260	0.4144	0.3235	0.3169	0.2772	0.3034	0.5479	0.4506	0.3024	0.2869
RWQMS	0.8171	0.8080	**0.9088**	**0.9077**	**0.8738**	**0.8728**	**0.8049**	**0.6873**	0.6386	0.6095
RRQA	**0.8798**	**0.8664**	0.8810	0.8715	**0.8465**	**0.8434**	0.6812	0.5291	0.7185	0.7112
CSQA	0.8732	0.8421	0.8538	0.8726	0.7723	0.7521	0.6523	0.6218	**0.7629**	**0.7712**
FQI	**0.8988**	**0.8863**	0.8903	**0.8907**	0.8124	0.8087	0.6923	0.6641	**0.8163**	**0.8152**

Table 4.12 : Performance comparison of FQI with respect to individual distortion type for QACS dataset. The top two performances are highlighted.

IQA Model	HVEC		HVEC-SCC	
	PLCC	SRCC	PLCC	SRCC
DNT	0.8200	0.8199	0.7804	0.7813
EPM	0.6540	0.6371	0.6504	0.6302
WNISM	0.6627	0.6423	0.5948	0.5479
FTB	0.6267	0.6266	0.7245	0.7376
SDM	0.6499	0.7395	0.6740	0.7205
RIQMC	0.5033	0.3382	0.4959	0.3354
RWQMS	**0.8478**	**0.8437**	**0.8407**	**0.8452**
RRQA	-	-	-	-
FQI	**0.9287**	**0.9211**	**0.9253**	**0.9289**

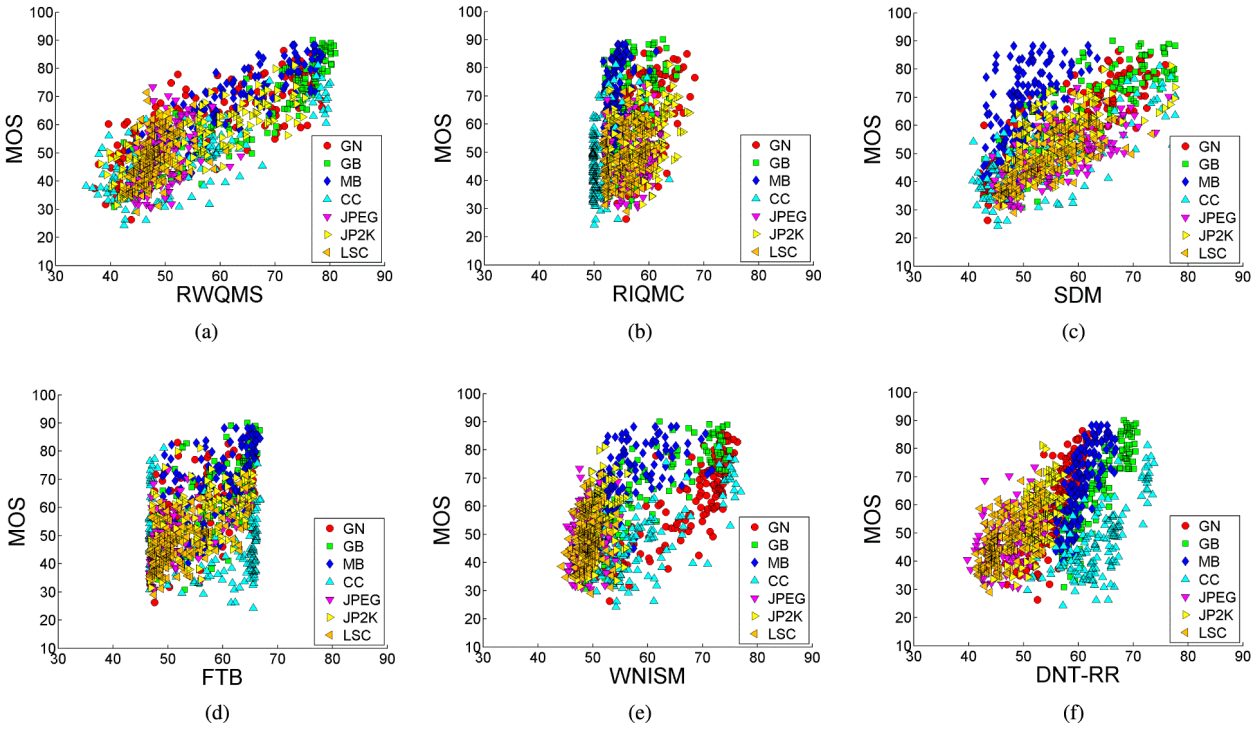


Figure 4.9 : Distortion wise Plot of Mean Opinion Score versus six state-of-the-art RR-IQA Methods on SIQAD dataset
(a) Reduced-reference wavelet-domain quality measure for SCIs (RWQMS) Wang *et al.* [2016], (b) reduced-reference image quality metric for contrast change (RIQMC) Gu *et al.* [2016c], (c) Structural degradation model (SDM), (d) Fourier transform-based scalable image quality measure (FTB) Narwaria *et al.* [2012], (e) Wavelet-domain natural image statistic model (WNISM) Wang and Simoncelli [2005],(f) Divisive normalization-based image representation (DNT-RR) Li and Wang [2009]

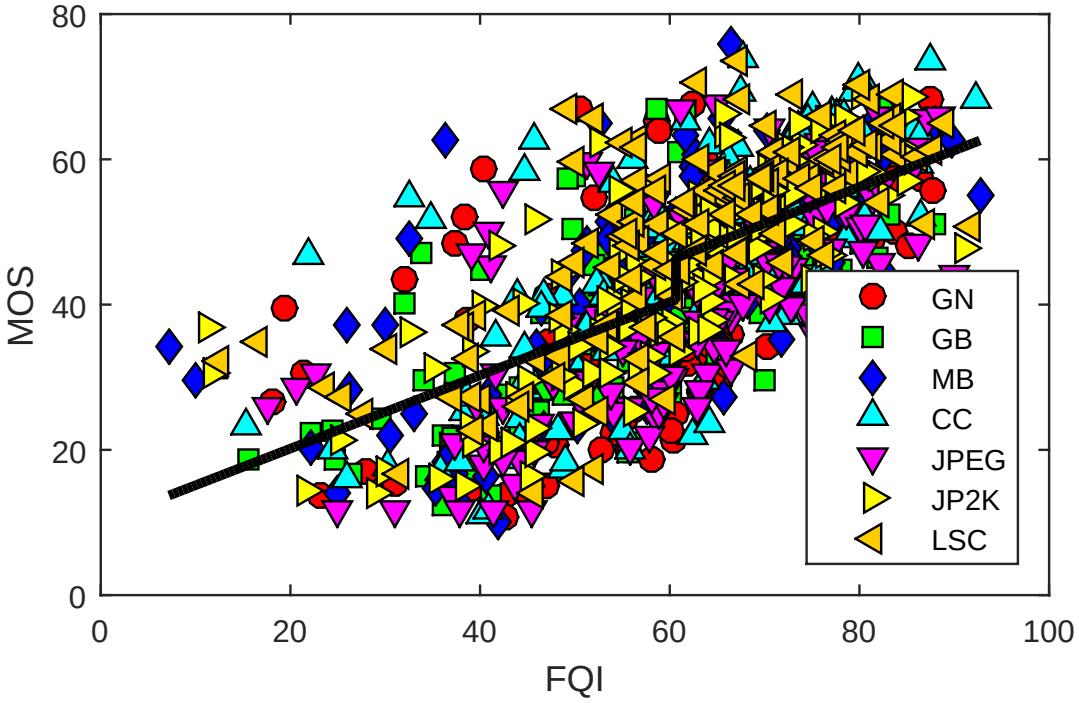


Figure 4.10 : Distortion wise Plot of Mean Opinion Score versus FQI on SIQAD dataset

result compared to RRQA Wang *et al.* [2018] under Gaussian noise and 12.5% under compression distortion. The reason for this performance improvement under distortion caused by compression can be the fact that the compression algorithms broadly try to preserve important features of the image. As the proposed method is feature based, it identifies any changes in the feature efficiently. Also, the pruned descriptors as mentioned in the FQI framework are more sensitive towards the distortion due to compression than contrast-change or motion blur as compression is an engineered distortion and contrast-change is a natural distortion.

The QACS dataset Wang *et al.* [2016] is focused only on the distortion caused by compression for SCIs. There are two types of distortions present in the dataset, such as HEVC, and HEVC-SCC which is the extension of HEVC for SCIs. The proposed method outperforms all other RR-IQA methods for both HEVC and HEVC-SCC as shown in Table 4.12. Although, as mentioned in Section 4.4.3, RWQMS Wang *et al.* [2016] which has been proposed along with the QACS dataset for quality assessment of distorted SCIs under compression distortion, appears as second best for the same dataset. FQI outperforms RWQMS with a margin of 9.5% for HEVC and 10.1% for HEVC-SCC. It can also be drawn from this result that FQI performs better for a compression algorithm specifically designed for SCIs.

For a better statistical comparison, the scatter plots of objective prediction for the six state-of-the-art RR-IQA methods in Li and Wang [2009]; Narwaria *et al.* [2012]; Gu *et al.* [2013, 2016c]; Wang *et al.* [2016]; Wang and Simoncelli [2005] and proposed FQI after regression versus subjective MOS on SIQAD dataset Yang *et al.* [2015] are shown in Fig. 4.9, and Fig. 4.10 respectively. The curves shown in Fig. 4.10 were obtained by a nonlinear fitting according to Larson and Chandler [2008]. From Fig. 4.9, and Fig. 4.10, it is observed that the objective scores predicted by proposed FQI correlate much more consistently with the subjective evaluations than the other state-of-the-art RR-IQA methods. Similar comparisons are portrayed for QACS dataset Wang *et al.*

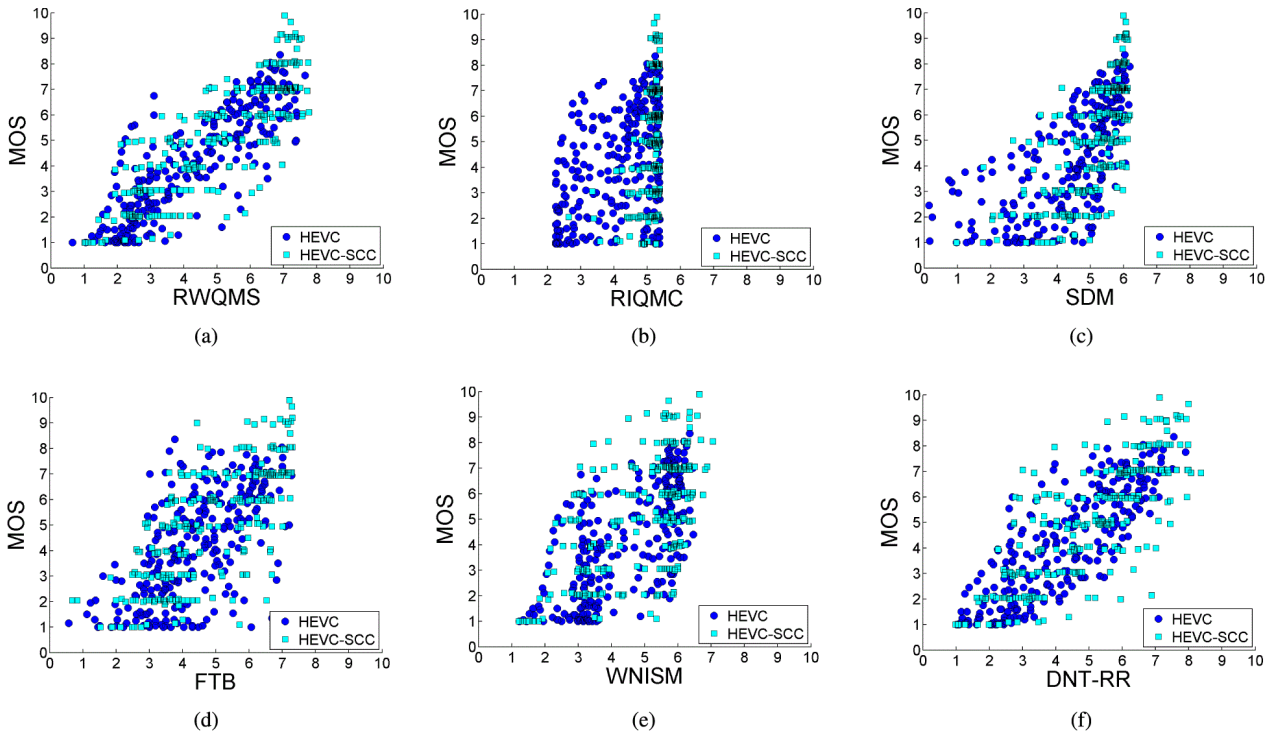


Figure 4.11 : Distortion wise Plot of Mean Opinion Score versus six state-of-the-art RR-IQA Methods on QACS dataset

(a) Reduced-reference wavelet-domain quality measure for SCIs (RWQMS) Wang *et al.* [2016], (b) reduced-reference image quality metric for contrast change (RIQMC) Gu *et al.* [2016c], (c) Structural degradation model (SDM), (d) Fourier transform-based scalable image quality measure (FTB) Narwaria *et al.* [2012], (e) Wavelet-domain natural image statistic model (WNISM) Wang and Simoncelli [2005],(f) Divisive normalization-based image representation (DNT-RR) Li and Wang [2009]

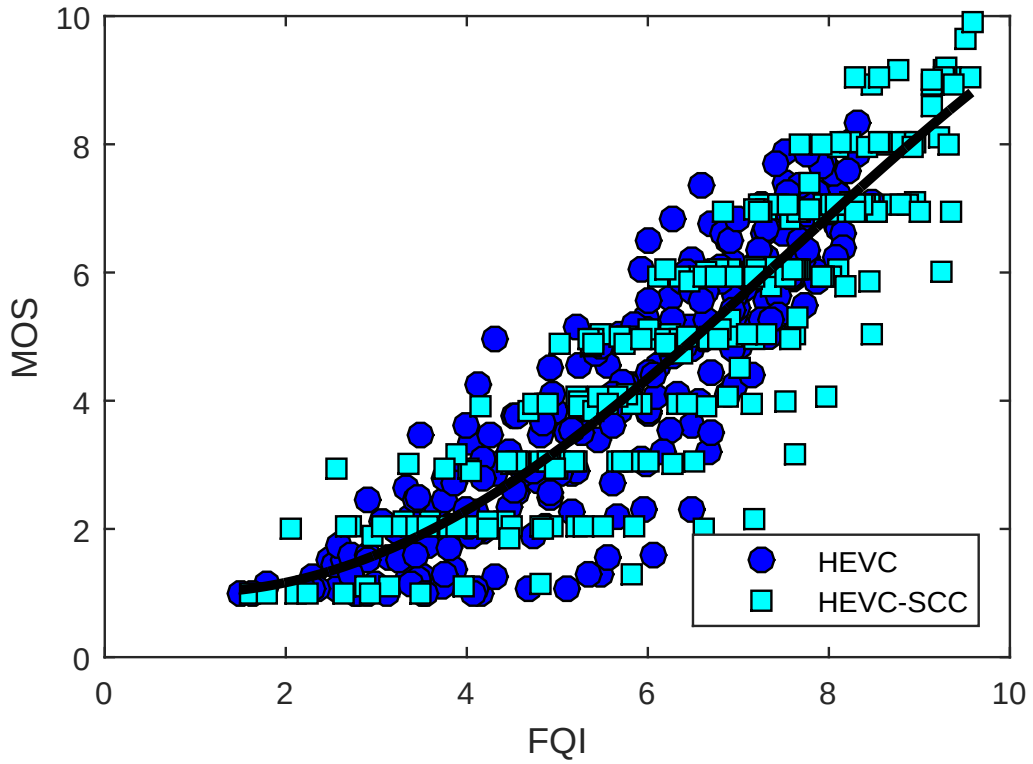


Figure 4.12 : Distortion wise Plot of Mean Opinion Score versus FQI on QACS dataset

[2016] in Fig. 4.11, and Fig. 4.12 to reflect the consistency of proposed FQI.

Table 4.13 : Computation cost analysis of proposed FQI

Feature-Matching Method	Features in RI	Descriptor Dimension	Features in DI	Descriptor Dimension	Distance Calculations	Total Computation
Conventional	$N1$	D	$N2$	D	$N1 \times N2$	$N1 \times N2 \times D$
SIFT	$N1$	128	$N2$	128	$N1 \times N2$	$N1 \times N2 \times 128$
CSQA	$N1$	128	$N2$	128	$N1 \times N2$	$N1 \times N2 \times 0.256$
FQI	$N1$	8	$N2$	8	$N1 \times N2 \times 0.002$	$N1 \times N2 \times 0.016$

4.4.5 Computation Cost Reduction in Proposed RDM

To provide certain invariance in the feature descriptors, the proposed feature extraction process came out to be somewhat computationally expensive. The proposed RDM for feature matching tries to recompense the computational complexity during feature extraction. The computational cost for RDM is validated on both SIQAD and QACS datasets. Typically, feature matching process for N1 features in the RI, and N2 features in the DI requires $(N1 \times N2)$ Euclidean distances to be computed, in order to compare the feature descriptors. The proposed feature matching process of FQI requires only 0.1% distance computations compared to the conventional feature matching process i.e. $(0.001 \times N1 \times N2)$. Moreover, the key-step in comparing two feature

points for a feature matching process is to evaluate the distance of every descriptor between the feature points. The proposed FQI yields only eight descriptors per feature and hence the computation cost for feature matching is further reduced. Table 4.13 illustrates the cost reduction under proposed FQI compared to traditional feature mating process where every feature has a *D* dimensional feature descriptors and SIFT feature matching Lowe [2004] where every feature has 128 dimensional descriptors.

4.5 CONCLUSIONS

Through this chapter, two feature-based RR-IQA methods for CCI and SCI are proposed. The underlying principle of the proposed CSQA and FQI is that the feature describes the perception of the human-visual-system (HVS) more than other parameters, such as intensity or structure. Also, as the feature of an image has a non-identical importance, it must be weighed for better quality assessment. CSQA outputs the quality metric in three steps: (i) extracting feature vector for reference image (RI) at sender side, and distorted image (DI) at receiver side, (ii) feature-matching to identify number of features preserved in the DI with respect to the RI, (iii) inner product of normalized scale and descriptor difference.

The feature matching of the proposed CSQA shows a significant reduction in computational cost of about (99.8%) compared to the traditional feature-based IQA techniques. The consistency, accuracy, stability, and robustness of the proposed CSQA are compared with 8 state-of-the-art RR-IQA methods, and 9 FR-IQA techniques for both CCI and SCIs. CSQA showed promising results for both CCI and SCI datasets. With about 0.21 bits-per-pixel overhead, the proposed CSQA outperformed all RR and FR IQA methods on QACS dataset and performed under top 3 for the rest of the datasets.

FQI produces the quality metric in four steps: (i) extracting feature vector for RI at the transmitter side, and DI at receiver side, (ii) feature-matching using reduced-distance method (RDM) to identify number of features preserved in the DI with-respect-to RI, (iii) normalization of scale value to weight the importance of the feature, and also normalization of descriptor distance of matched features, (iv) inner product of normalized scale and descriptor difference.

With a reduced set of feature descriptor in FQI compared to CSQA, its feature extraction method provides invariance towards scaling and rotation which are helpful during feature matching. Also, the feature matching process of the proposed FQI shows a significant computational cost reduction (99.8%) compared to the traditional feature-matching techniques where to find suitable feature match, every feature point in RI is compared with all the feature points in DI. The consistency, accuracy, and robustness of the proposed FQI are compared with 9 state-of-the-art FR-IQA, and 8 RR-IQA techniques. FQI has shown promising results as it outperformed all RR and FR-IQA methods for QACS dataset, and performed as second best for SIQAD dataset. The comparison of individual distortion type shows that the proposed FQI is more suitable towards distortion caused by compression, and it can be proven very useful for such research communities.

…