A novel image quality assessment method and coefficient of quality for digital solutions of colour blindness

Meenakshi S | Anshu Singla

Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India

Correspondence
Anshu Singla, Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India.
Email: anshu.singla@chitkara.edu.in

Abstract
Eyesight is one of the primary senses that human beings have. Reports show that colour blindness, a form of colour vision deficiency (CVD), affects about 8% of the male population and 0.5% of female population. The Assistive Technology Act of 2004 lays focus on technologies that help individuals with disabilities and deficiencies. With the rapid advancement in technologies, several assistive solutions are available for visually impaired or CVD patients. Such solutions involve simulation and compensation of conflicting colours to help the colour blind in the visual perception of colours. Given the increased usage of the web, post the pandemic, these solutions improve the quality of life for the colour blind. Defining the image quality assessment criteria for such digital solutions becomes imperative. The study proposes a novel method for image quality assessment of digital solutions aimed at assisting the colour blind users. The proposed coefficient of quality ($CQ$) would be useful to rank colour compensation and recolouring algorithms. Experiments were conducted with a novel questionnaire set designed for this quality measurement. The results affirm the efficiency of the assessment method proposed. This will also provide objective feedback to the researchers and experts in this area to improve their solutions for CVD patients.

1 | INTRODUCTION

With the rapid development and rampant use of web for daily routine activities, there has been an increased awareness on making it accessible to people with vision deficiencies [1]. Colour blindness is one form of colour vision deficiency (CVD) which affects about 8% of the male population and 0.5% of female population [2]. Due to this deficiency, the patients suffering from this anomaly, referred to as colour blind are unable to distinguish colours significantly. In order to assist the colour blind for viewing digital images on the web, a number of colour compensation/correction methods have been developed [3]. These colour compensation methods perform colour correction on the conflicting colours thereby improving the perception of them by the colour blind. Efficiency of the colour compensation methods is generally measured through subjective evaluation carried out by colour blind and/or normal vision volunteers. The parameters used for such subjective evaluation is a challenge as human observers (evaluators) are involved. Hence the criteria need to be defined scientifically to rank and measure the efficiency of the colour compensation methods. Such criterion should measure the effectiveness of the method as to how it maintains image quality while recolouring, at the same time improve the perception of the conflicting colours for the colour blind. In this section, the terminology used in the area of CVD, its solutions and parameters for image quality assessment are defined.

1.1 | Colour vision deficiency

Colour vision deficiency (CVD) refers to the inability to distinguish certain colours from other colours [4–5]. The photoreceptors known as rods and cones in human eyes are responsible for colour vision [6–7]. Rods are responsible for peripheral vision and cones are responsible for colour vision. There are three classes of cones namely L, M and S corresponding to the three wavelengths: (i) Long-wavelength (L),...
(ii) Middle-wavelength (M), and (iii) Short-wavelength (S) regions of the spectrum. Any stimulus for colour is specified by the LMS values. The normal colour vision is trichromatic and happens due to the absorption of photon by all the three cones present in the eyes. Anomaly in colour vision is primarily caused by defects or mutations in the genes to make photo pigments present in the eyes. Anomaly in colour vision happens due to the absorption of photon by all the three cones regions of the spectrum. Any stimulus for colour is specified by the LMS values. The normal colour vision is trichromatic and happens due to the absorption of photon by all the three cones present in the eyes. Anomaly in colour vision is primarily caused by defects or mutations in the genes to make photo pigments present in the eyes. Anomaly in colour vision happens due to the absorption of photon by all the three cones regions of the spectrum. Any stimulus for colour is specified by the LMS values. The normal colour vision is trichromatic and happens due to the absorption of photon by all the three cones present in the eyes. Anomaly in colour vision is primarily caused by defects or mutations in the genes to make photo pigments present in the eyes. Anomaly in colour vision happens due to the absorption of photon by all the three cones regions of the spectrum. Any stimulus for colour is specified by the LMS values. The normal colour vision is trichromatic and happens due to the absorption of photon by all the three cones present in the eyes. Anomaly in colour vision is primarily caused by defects or mutations in the genes to make photo pigments present in the eyes. Anomaly in colour vision happens due to the absorption of photon by all the three cones regions of the spectrum. Any stimulus for colour is specified by the LMS values. The normal colour vision is trichromatic and happens due to the absorption of photon by all the three cones present in the eyes. Anomaly in colour vision is primarily caused by defects or mutations in the genes to make photo pigments present in the eyes. Anomaly in colour vision happens due to the absorption of photon by all the three cones regions of the spectrum. Any stimulus for colour is specified by the LMS values. The normal colour vision is trichromatic and happens due to the absorption of photon by all the three cones present in the eyes. Anomaly in colour vision is primarily caused by defects or mutations in the genes to make photo pigments present in the eyes. Anomaly in colour vision happens due to the absorption of photon by all the three cones regions of the spectrum. Any stimulus for colour is specified by the LMS values. The normal colour vision is trichromatic and happens due to the absorption of photon by all the three cones present in the eyes. Anomaly in colour vision is primarily caused by defects or mutations in the genes to make photo pigments present in the eyes. Anomaly in colour vision happens due to the absorption of photon by all the three cones regions of the spectrum. Any stimulus for colour is specified by the LMS values. The normal colour vision is trichromatic and happens due to the absorption of photon by all the three cones present in the eyes. Anomaly in colour vision is primarily caused by defects or mutations in the genes to make photo pigments present in the eyes. Anomaly in colour vision happens due to the absorption of photon by all the three cones regions of the spectrum. Any stimulus for colour is specified by the LMS values. The normal colour vision is trichromatic and happens due to the absorption of photon by all the three cones present in the eyes. Anomaly in colour vision is primarily caused by defects or mutations in the genes to make photo pigments present in the eyes. Anomaly in colour vision happens due to the absorption of photon by all the three cones regions of the spectrum. Any stimulus for colour is specified by the LMS values. The normal colour vision is trichromatic and happens due to the absorption of photon by all the three cones present in the eyes. Anomaly in colour vision is primarily caused by defects or mutations in the genes to make photo pigments present in the eyes. Anomaly in colour vision happens due to the absorption of photon by all the three cones regions of the spectrum. Any stimulus for colour is specified by the LMS values. The normal colour vision is trichromatic and happens due to the absorption of photon by all the three cones present in the eyes. Anomaly in colour vision is primarily caused by defects or mutations in the genes to make photo pigments present in the eyes. Anomaly in colour vision happens due to the absorption of photon by all the three cones regions of the spectrum.
Table 1 shows the image as viewed by the colour blind and how the colour compensation can help.

When images are colour compensated, it becomes necessary to provide methods to check the efficiency of such enhancements from both a colour blind and image quality perspective. The next section provides the details of image quality assessment methods.

### 1.3 Image quality assessment methods

Image quality assessment (IQA) methods measure the level of quality, specification and condition of being precise. IQA is a part of the quality of experience (QoE) measures. These methods are employed to monitor image quality such that image processing algorithms and systems are optimized. IQA methods can be further broken down into two categories: (i) Subjective image quality assessment methods that involve the subjective judgement of the number of observers on the basis of which the rating is done on an absolute scale—excellent, good, fair, poor, or bad. The detailed stratification of this method is shown in Figure 2(a). (ii) Objective image quality assessment methods that involve a mathematical model to predict the quality of image accurately and quantitatively. The stratification of objective IQA methods is shown in Figure 2(b).

The objective of these methods is to emulate the predictions of the evaluators.

### 2 REVIEW OF EXISTING LITERATURE

Evaluators who may be colour blind or normal vision users evaluate the effectiveness of colour compensation on images proposed by authors. There are two aspects to this effectiveness evaluation. One is the effectiveness of perception of the conflicting colours by the colour blind in the compensated image. The second is the image quality aspects following the compensation. Hence the parameters for image quality assessment for the colour compensation methods in case of colour blind must contain parameters from both these aspects in a balanced manner. Evaluating image quality perspective alone is not sufficient although the International Telcom Union (ITU) has set a number of standards for subjective quality assessments of television pictures, digital video quality and the brightness and contrast of display devices [20–23].

Pedram Mohammadi et al. have reviewed the quality assessment methods for images including six popular subjective quality datasets [24]. They propose that the trend is towards automated quality assessments in the area of medicine and other allied sciences [24]. X. Yang et al. cover various deep neural network methods for BIQA. They also compare the performance of these methods against the synthetic databases available [25]. Athar et al. have surveyed the image quality assessment...
algorithms and compared their performance [26]. They propose rank aggregation based training-free FR fusion methods as they would provide a robust perceptual quality prediction performance when validated across subject-rated datasets. Aravind Babu Jeriopothula et al. have created an approach for assessing quality of images using the BIQA approach motivated by three techniques applicable for natural images [27]. They propose two shallow CNN models coupled with enriched Natural Scene Statistics (NSS) feature which works well on natural images. A quality score is also proposed. Zhai G et al. have surveyed the state-of-the-art methods and classical methods in the area of perceptual image quality assessment and proposed that the trend is towards stereoscopic images and saliency guided approaches [28]. Lihuo He et al. have surveyed the objective IQA measures and proposed a computational model to predict human perception of image quality. The recent day metrics focus on natural image quality assessment (NIQA) [29]. This involves having a good reference image database, knowledge on the human visual system and distortion parameters. They compare several methods on these perspectives. The literature studied suggests a movement towards automated quality assessment of images through deep learning techniques and without a reference image (BIQA). Blind image quality assessment (BIQA) methods using deep learning methods are being used to envisage the image quality as to how human beings will perceive without the need for a reference image [30]. There is focus on maintenance of naturalness in the image in addition to maintenance of image quality like contrast, distortion ratio etc. However, with respect to our area of study, IQA methods suggested are generic in nature and do not specifically target the colour blindness scenario. BIQA techniques may not be directly applied as image databases specific to colour blindness of varying degrees is not available. Also the image databases should be able to simulate natural scenarios and have a mix of conflicting which is difficult to obtain in the sample databases. However, parameters and metrics like Mean significant ratio (MSR) and sum of the squares of the deviations (SSD) can be utilized while defining parameters for assessing image quality of compensation for colour blind. Also, the parameters like maintenance of brightness, intensity and contrast can be considered while evaluating the effectiveness of compensation.

In addition, the authors also studied the state-of-the-art methods that have been used to evaluate of effectiveness of colour compensation algorithms. The observation shows that subjective feedback is elicited on the compensated images from the colour blind or simulation is done on the compensated images and feedback from normal vision users is sought [31–35]. Other authors have compared the compensated images with the state-of-the-art and checked the effectiveness in terms of contrast, clarity and computation time [36–40]. A couple of authors have used Thurstone’s Law of Comparative Judgment for seeking subjective feedback from the colour blind on the compensated images [41–42]. In 2017, Simon-Liedtke et al. provided guidelines for the design and evaluation of algorithms using daltonization method for colour compensation [43]. The details of the different state-of-art techniques existing in literature have been categorized in two types: (a) Subjective (SUB) that is subjective evaluation by evaluators and (b) Objective (OBJ) that is evaluation of algorithm through comparison with state-of-the-art or simulated view of how colour blind would see. The analysis is given in Table 2.

![Table 2](image)

The literature studied yields the following limitations: (i) Measurement of the image quality post the compensation to evaluate the effectiveness has been done by using subjective feedback from normal vision and colour blind without using standard scales for the responses; (ii) this feedback does not cover image quality parameters and is based on the overall perception of the evaluator; (iii) when an image is compared with another image (test image vs. reference image), the outcome is not measured using objective image quality assessment methods. The authors present a research work which is a novel subjective image quality assessment method for evaluation of the digital solutions that perform colour compensation of images for the colour blind and a benchmark metric that will quantify the evaluation called the coefficient of quality \( C_Q \). The proposed method has two parts. Part I is a subjective image quality assessment method to be used when there are no reference images available and the evaluation is done through subjective evaluation by evaluators (colour blind or normal vision people). Part II is an objective image quality assessment which involves the use of statistical measures that compare the effectiveness of the digital solution against the state-of-the-art methods. This is used when there is lack of colour blind evaluators/ CVD patients to evaluate the solutions but a reference image is available.

The key research contributions by the authors in this study are: (a) a questionnaire covering the parameters from a colour blind and image quality perspectives; (b) method to arrive at the benchmark metric \( C_Q \); (c) identification of the state-of-the-art statistical metrics for objective quality assessment. The \( C_Q \) would be a good benchmark on quality perception in
image processing specifically in the area of colour compensation/correction/simulation algorithms for CVD. The authors believe that it can be used to rank the various algorithms and methods for CVD and provide valuable feedback on the specific improvement areas to the research community as the questionnaire is prepared with all perspectives given by optometrists and image analysis experts.

This paper is organized as follows: the details of the method and computation of the metric $CQ$ is described in Section 3. The experiment conducted to test the method, results and the analysis is discussed in Section 4. Section 5 provides the further improvements and Section 6 provides the conclusion for this research work.

3 | PROPOSED IMAGE QUALITY ASSESSMENT METHODS

Digital solutions involving colour compensation methods acquire the image, segment out the conflicting colours as per the type of colour blindness and recolour those colours for better perception of the colour blind. In this process, the image quality perspective is the first factor that should be maintained. The other factor is the conflicting colour should be perceived by the colour blind without affecting the other colours and naturalness of the image. The authors propose a novel subjective assessment method and a benchmark metric $CQ$ taking into account both the factors. The value of this metric can be used for ranking the effectiveness of the colour compensation algorithms specifically for CVD users. The authors propose statistical measures as part of the objective quality assessment method to be used in cases, where the reference image is available.

### 3.1 Proposed novel subjective image quality assessment method and benchmark metric $CQ$ for image quality assessment

When there is subjective evaluation to be done on a range of questions by evaluators, the scoring can be done using Thurstone scale [44]. If only one image is used, then it is categorized as single stimulus method as depicted in Figure 2(a). The steps involved in proposed novel subjective quality assessment method are shown in Figure 3. The method has been designed in alignment with all key factors as specified in Section 2.

Step 1: Develop a series of questions. This step involves arriving at questions on which the opinions need to be measured for evaluating the colour compensation. The questionnaire developed comprise 19 questions based on the image quality parameters [21–23]: (i) naturalness in preserving, (ii) preservation of optimal colours, (iii) recolouring, (iv) type of image, (v) CVD related and (vi) testing as illustrated in Figure 4. The questions designed cover the parameters like naturalness preserving, preservation of optimal colours, image quality parameters like intensity, brightness and the testing being done. For example, the preservation of brightness and intensity of the colours constitute naturalness preserving parameters. The colour compensated image should retain the mood of the original image which constitutes the preservation of optimal colours. Intensity and brightness of conflicting and remaining colours in the image are measures of image quality. Ideally, the colour compensated images need to be tested with real CVD patients and on various types of images which connotes the testing parameters. The possible responses to the questions are also mentioned in the questionnaire as shown in Figure 4.


**TABLE 3** Mapping of scores to the response values

| Response | Score ($S_r$) |
|----------|--------------|
| Excellent | 5            |
| Good     | 4            |
| Fair     | 3            |
| Poor     | 2            |
| Bad      | 1            |
| Yes      | 1            |
| No       | 0            |

Step 2: Assign a weighted score to the questions. While the questionnaire would cover the factors mentioned in the form of parameters, all of them may not be of equal importance. Some parameters may be of more importance than the others. For example, in case of visual perception for the colour blind, the perception of the conflicting colour is more important compared to the brightness of the colour. In statistics, a weight function allows to allocate more importance or weight to some elements in a set [45]. They are used on data that is measured for both discrete and continuous types. Hence, each question created is given a weight score on a scale of 1 to 11 based on the importance of the parameter to be fulfilled. This weightage ($Q_w$) has to be given by a panel of judges in the area to avoid author bias in terms of inclination to a certain parameter or technique and provide an expert opinion on the parameters. For example, as an author of a colour compensation algorithm, the focus may be on ensuring that the conflicting colour is recoloured (author bias). However from the colour blind perspective, an optometrist would want to focus on recolouring with retention of the objective and mood of the image that is preserve naturalness and also colour compensate based on the degree of the colour blindness.

Step 3: Calculate median/interquartile range ($IQR$). Calculate the $IQR$ value for the weightage scores given by the panel of experts for the questions. The $IQR$ describes the middle 50% of values when sorted from lowest to highest [46]. $IQR$ is helpful in identifying whether the value is an outlier. It also helps identify if the outlier is mild or strong. In the proposed method, the $IQR$ values can help in filtering out questions if a smaller set is needed by removing the outliers. Questions with extreme values can be removed. The questionnaire developed had 19 questions. The questions, factor considered, weightage assigned and the $IQR$ value are shown in Figure 4.

Step 4: Record responses by evaluators. The next step is to provide this questionnaire to evaluators. In our study, the evaluators are colour blind or normal vision human observers. There are two scales that can be used by evaluators for recording the responses in the questionnaire: nominal scales (Yes/No) and continuous scale with values excellent, good, fair, poor and bad. Each of the responses are then mapped to a numerical score value as shown in Table 3.

To record responses, in the experiment, nominal scales are used for question numbers 4–6 and 17. Continuous scales used for question numbers 1–2, 7–11, 13–16 and 18 are shown in Figure 4.

Step 5: Computation of the proposed coefficient of quality ($CQ$): In this step, the subjective responses are converted to a quality coefficient using the following statistical Equations (11)–(14). A weighted mean is also a kind of average; however, it is used when the all the elements do not contribute equally [46]. The weighted mean score for each question ($Q_w$) is calculated as given in Equation (11). $J_w$ is the weightage given by the judge for each question where $m$ represents the number of judges.

$$Q_w = \sum_{x=1}^{m} J_w$$

(11)

Let the number of evaluators be $n$. The weighted mean response weightage score $R_w$ for each question is calculated as given in Equation (12). $S_r$ is the score value given to the response given by the evaluator to a question.

$$R_w = \sum_{x=1}^{n} Q_w \times S_r$$

(12)

The weighted mean ideal weightage score $I_w$ is calculated as given in Equation (13). $S_i$ is the ideal score for every question. For the nominal scale it is 1 and for continuous scale it is 5.

$$I_w = \sum_{x=1}^{n} Q_w \times S_i$$

(13)

Let the number of questions in the questionnaire be $q$. $CQ$ is calculated as given in Equation (14).

$$CQ = \frac{\sum_{x=1}^{q} R_w}{\sum_{x=1}^{q} I_w}$$

(14)

Table 4 shows sample responses of an evaluator on the questionnaire for image 5 in Figure 5 and the computation of $CQ$.
value using the proposed method.

In a number of colour compensation methods for colour blindness, reference images are available which have predefined colours and objects that can be perceived by normal vision users. In such cases, our proposed method would involve the steps shown in Figure 5.

Step I: Record responses of colour blind evaluators on reference image. Evaluators observe the reference image and record their observations on what they see.

Step II: Record responses of colour blind evaluators on colour compensated image. Evaluators observe the colour compensated image and record their observations on what they see.

Step III: Score the responses. Score the responses based on whether the colour blind evaluator is able to identify the colours and objects in the reference and colour compensated images. A sample scoring is shown in Table 5.

Step IV: Computation of the proposed $CQ$. In this step, the response scores for the reference and colour compensated images are compared. The score comparison is converted to a quality coefficient which provides a measure of the effectiveness of the colour compensation. The mathematical definition for the steps using weighted means is given in this section.

Let the number of evaluators (colour blind) be $n$. The images being evaluated by the evaluators may have a weightage associated based on the components being present in the image. An image which is viewed by both normal vision and colour blind people may be provided to ensure that the evaluation is accurate. Hence each of the images may be associated with weightage as well. Weighted mean response weightage score is calculated for the reference image ($R_w$) and for the colour compensated image ($R'_w$) as given in Equations (15) and (16). $P_w$ is the weightage given to the image, $S_r$ and $S'_r$ are the score values against the responses provided by the evaluators for reference and colour compensated image, respectively.

$$R_w = \sum_{x=1}^{n} P_w \times S_r$$  \hspace{1cm} (15)

$$R'_w = \sum_{x=1}^{n} P_w \times S'_r$$  \hspace{1cm} (16)

The ideal weightage is $I_w$ calculated as given in Equation (17).

$$R'_w = \sum_{x=1}^{m} P_w$$  \hspace{1cm} (17)

$CQ$ is calculated as given in Equation (18).

$$CQ = \frac{(R'_w - R_w)}{I_w}$$  \hspace{1cm} (18)

Tables 6(a) and 6(b) show sample responses of an evaluator to five images (reference and colour compensated) and the computation of $CQ$ value using the proposed method.

In this subjective evaluation method, inferences can be derived from the coefficient value and the closeness to 1 indicates good efficiency of recolouring. The weightages are derived from views of experts in the vision and image processing area in case of the single stimulus method, the possibility of bias is greatly reduced. Also, in this method, the subjective evaluation is not based on an overall perception of the evaluator, but on specific image quality and aspects from the perspective of improved visualization for the colour blind users. In case of the double stimulus method, the evaluation is against the

| Image no. | Normal vision value | Weightage for the images | Responses from evaluator for reference image | Score for the responses ($S_r$) |
|-----------|---------------------|--------------------------|---------------------------------------------|-------------------------------|
| 1         | 12                  | 1                        | 12                                          | 1                             |
| 2         | 8                   | 3                        | 3                                           | 0                             |
| 3         | 29                  | 3                        | 70                                          | 0                             |
| 4         | 5                   | 3                        | Do not see anything                        | 0                             |
| 5         | 3                   | 3                        | 5                                           | 0                             |

$I_w = 13$
TABLE 6b  \( CQ \) computation for subjective image quality assessment  
(double stimulus)

| Image no. | Responses-colour compensated image | \( S'_i \) | \( R_w \) | \( R'_w \) |
|-----------|-------------------------------------|---------|-------|-------|
| 1         | 12                                  | 1       | 1     | 1     |
| 2         | 3                                   | 0       | 0     | 0     |
| 3         | 20                                  | 1       | 0     | 3     |
| 4         | 5                                   | 1       | 0     | 3     |
| 5         | 3                                   | 1       | 0     | 3     |
| Total     |                                     | 1       | 10    |       |

\( CQ = (10-1)/13 = 0.69 \)

reference and colour compensated images. The scores on the questions obtained can provide specific areas of improvement for the authors of the algorithm. The \( CQ \) metric can be used to rank various colour compensation algorithms for colour blindness as well.

3.2  | Proposed statistical measures of objective image quality assessment

Objective image quality assessment does not rely on subjective feedback, rather employs automated and statistical measures to obtain the image quality based on image quality metrics [5]. It is a common challenge to identify a sizeable sample colour blind population for observations. In such cases, the evaluation of the colour compensation can be achieved by simulating the colour compensated image as to how a colour blind would see. Thus simulation methods of how a colour blind would see plays an important role in image quality assessment. To measure the effectiveness of the simulation technique, a comparison is done with standard simulation tools or state-of-the-art. The simulated image by the author of the solution is compared with the simulation generated by the standard tool for similarity. Standard statistical image comparison measures can be used to do this comparison for effectiveness of simulation. The higher the similarity, the better is the effectiveness of the simulation technique. The authors propose using the following statistical measures as part of objective measurement of colour blind simulation methods.

I. Pixel based comparison of images: A pixel-by-pixel comparison is useful when the images being considered for comparison have identical characteristics and conditions (size, resolution, type etc.)

II. Comparison using scale-invariant feature transform (SIFT) algorithm [47]: SIFT is invariance to image scale and rotation and is a patented algorithm. It allows for feature detection in computer vision applications. This can be used for comparison of two images. It extracts the features from the reference image and stores in a database. The features of the test image are then compared with the database using candidate matching features based on Euclidean distance of the respective vectors. The more the matches, the higher the accuracy is.

III. Sum of squared differences (SSD) between the matching points in the given images: SSD is a measure of match between two images and calculates the summation of the squared for the product of pixels subtraction between the two images.

A combination of a basic image comparison measure (pixel comparison), feature extraction measure (SIFT) and a template matching measure (SSD) to give an accurate view of the image comparisons.

4  | EXPERIMENTAL ANALYSIS  
AND RESULTS

This section presents the experiment conducted to test the proposed methods and the analysis of the results. It also mentions the data set used and describes the volunteers for the evaluation.

4.1  | Benchmark data set and volunteers

The data sets used in this study are 24 Ishihara plates [18]. The images in bitmap of 24 bit have been chosen as shown in column 2 (Original image) of Figure 6. The images have been carefully chosen with conflicting colour backgrounds to validate the image quality assessment indicating the effectiveness of the colour compensation for all types of colour blindness. For example, for the sample images shown in Figure 6, in the image 4, pink and yellow are present in the flower. Yellow is perceived as pink by tritanopes. In image 5, blue and green colours are present in the image. Tritanopes perceive blue as green.

There were two sets of evaluators. The first set consisted of 10 normal vision volunteers in the age group of 10–40. These
evaluators participated in the proposed subjective image quality assessment (single stimulus) method. The second set consisted of 3 colour blind evaluators who participated in the subjective image quality assessment (double stimulus) method. Colour blind evaluator 1 suffered from protanopia, colour blind evaluator two suffered from deuteranopia and colour blind evaluator three had a combination of red-green colour blindness (protan-deuteran). The images were shown to volunteers using the guidelines mentioned by ITU [21–23] with appropriate monitors, spacing and time gap between image observations. Ten optometrists and image processing experts participated in providing the weightages for the questions based on their expertise in working with colour blind patients and image processing areas, respectively.

The recolouring solution was also chosen to represent a case of effective colour compensation and that which has not yielded a relatively effective compensation to demonstrate the quantitative nature of the benchmark metric $C_Q$.

### 4.2 Experiment and results for proposed novel subjective image quality assessment method

The authors took the image sets [48] and performed the colour compensation using the method present in Section 1.2. The sample results are shown in Figure 5.

The compensated colour and simulated images were shown to the 10 normal vision evaluators. They evaluated the effectiveness of the colour compensation using the questionnaire given in Figure 4 using the responses mentioned in Table 3. Two questions from the questionnaire shown in Figure 4, with question number 3 and 12 were not used in this experiment as they were pertaining to maps. A sample response from one of the evaluators is shown in Table 7.

The $C_Q$ values computed for five images are provided in Table 7. The computation is done by following the mathematical structure provided in step 5 of Section 3.1 (single stimulus).

The $C_Q$ is highest for image 5 and lowest for image 3 indicating that the colour compensation has been relatively effective in image 5 when compared to the rest of the images. The responses for image 5 given in Table 4 indicate that the algorithm needs to focus on (a) Testing on real CVD patients (b) address all forms of colour blindness (c) focus on naturalness preserving aspects (d) test it on black and white images too. Thus these are specific feedback to author of colour compensation methods for improvement on image quality parameters. This means the questionnaire formed informs the subject experts, authors, and researchers about the aspects that need to consider improving the algorithm.

The authors also performed the experiment to test the double stimulus method. The colour blind evaluators were shown the original Ishihara plates and what they visualize in the plates was recorded. Then the colour compensated Ishihara plates using the method explained in Section 1.2 were shown to the colour blind evaluators and what they see in this case was also recorded. The scoring was done as provided in Table 5 based on whether the colour blind evaluator was able to make out the pattern present in the Ishihara plates fully, partially or could not see. This was done for both the reference Ishihara and colour compensated Ishihara plates. The $C_Q$ values were computed as per the mathematical definition given in step IV of Section 3.1. The results are shown in Table 8.

The $C_Q$ values for the Ishihara plates indicate that the colour compensation is relatively better for red-green colour blindness (as per the result from colour blind evaluator 3) while it is relatively poor for protanopia and deuteranopia, respectively (as per result from colour blind evaluator 1 and 2). This also shows that the colour compensation method used for Ishihara plates is not very effective.

Thus in both these experiments and results, the proposed novel subjective image quality assessment method quantitatively measures the effectiveness of the colour compensation and provides specific areas of improvement for the author or researcher of the colour compensation methods.

### 4.3 Experiment and results for proposed statistical measures of objective image quality assessment

To measure the effectiveness of colour blind simulation technique, the authors used the images in the mentioned data set [48]. Simulated images were generated using the method explained in Section 1.2. To compare the effectiveness of the mentioned simulation, Gobis [49], a state-of-the-art colour blind simulation tool was used and the simulated images were generated as shown in Figure 7(a). The images were correspondingly compared using the statistical measures: (i) Pixel based comparison (ii) SIFT, and (iii) SSD as mentioned in Section 3.2, the results are shown in Table 9.

The results indicate that the similarity is maximum for image 1 and least for image 2. It can be observed that the average difference 2.90% for the five images considered implying an

| Image no. | $C_Q$ |
|-----------|-------|
| 1         | 0.85  |
| 2         | 0.74  |
| 3         | 0.63  |
| 4         | 0.75  |
| 5         | 0.88  |

| Evaluator no. | $C_Q$ |
|---------------|-------|
| 1 (Protan)    | 0.32  |
| 2 (Deuteran)  | 0.33  |
| 3 (Protan-deuteran) | 0.40 |
TABLE 9  Results of pixel based comparison for colour blind simulation

| Image no. | Total no. of pixels | Difference in pixels | Difference percentage | SSD values |
|-----------|---------------------|----------------------|-----------------------|------------|
| 1         | 12804               | 95                   | 0.74%                 | 1.30E-08   |
| 2         | 47808               | 3265                 | 6.83%                 | 1.62E-06   |
| 3         | 27456               | 1127                 | 4.11%                 | 3.70E-06   |
| 4         | 27456               | 508                  | 1.85%                 | 2.84E-05   |
| 5         | 27456               | 270                  | 0.98%                 | 1.38E-06   |
| Average   | 142980              | 5265                 | 2.90%                 | 7.02E-06   |

TABLE 10  Comparison of proposed benchmark metric with state-of-the-art methods

| Author’s method               | Effectiveness metric as provided by the authors (%) | Effectiveness by proposed benchmark metric (%) |
|-------------------------------|-----------------------------------------------------|-----------------------------------------------|
| Pinheiro Dos Santos W’s method [13] | 47%                                                  | 50%                                           |
| Zhenyang Zhu’s method [50]    | 81%                                                  | 82%                                           |
| Li’s method [51]              | 61%                                                  | 69%                                           |

accuracy of 97.1%. The SSD values indicate that the similarity is maximum for image 1 and least for image 4 shown in Figure 5.

The comparison of the images mentioned was also done by using SIFT method. The results of applying this method for one of the images are shown in Figure 7(a) and 7(b). Figure 7(a) shows the epipolar line sketch and Figure 7(b) indicates the points which were matched.

It can be observed that there is similarity in the epipolar lines and a number of points are also matched between the images indicating high degree of similarity. This indicates that the colour blind simulation technique is effective in a quantitative manner.

4.4  Comparison of results of proposed subjective assessment methods with state-of-the-art

In order to validate the proposed subjective, objective image quality assessment method and the benchmark metric $C_Q$, the authors studied the colour compensation effectiveness measure-
TABLE 11 Comparison of results of proposed objective assessment methods with state-of-the-art

| Author's method                  | Effectiveness metric as provided by the authors | Effectiveness by proposed statistical measure (pixel comparison) | SSD values |
|----------------------------------|-------------------------------------------------|---------------------------------------------------------------|------------|
| Zhenyang Zhu's method [50]       | 0.99 (Average chromatic difference for protanopia) | 0.95                                                          | 6.08E-07   |
| Simon-Liedtke's method [52]      | 0.94 (z score)                                  | 0.90                                                          | 4.01E-07   |

FIGURE 8 Point matching using SIFT method for I1 and I3 for [50]

respectively). The authors generated the pixel based statistical comparison to find out the similarity between I1 and I3, and also for I2 and I4. This statistical measure was compared with the statistical measure proposed by [50, 52]. The results are shown in Table 11. The matching points obtained by applying the feature extraction algorithm SIFT on I1 and I3 are shown in Figure 8 for [50]. The matching points obtained by applying the feature extraction algorithm SIFT on I2 and I4 are shown in Figure 9 for [52].

The results in column 2 of Table 11 indicate that the statistical measures proposed by the authors namely pixel comparison are very close to those measured by the state-of-the-art literature [50, 52]. The SSD values given in column 3 of Table 11 are very minimal indicating a high degree of similarity (0.99 similarity). This is similar to the results given by the authors in [50] and [52] as shown in column 2 of Table 11. The number of points matched between I1 and I3 for [50] and I2 and I4 for [52] as shown in Figures 8 and 9, respectively, also indicate a high degree of similarity. This validates that the proposed statistical measures are effective.

Additionally, the three proposed statistical measures constitute a combination of basic image comparison method (pixel differences), a feature extraction method which provides good accuracy in image comparison (SIFT) and a template matching method (SSD) would provide a good measure of image comparisons than using only individual statistical methods. Hence the combined statistical measures proposed connote a good measure of the objective image quality assessment. Considering all the results discussed, it can be concluded that the simulation using the proposed subjective and objective image quality assessment method has good accuracy.

In this study, the proposed method for subjective image quality measurement has yielded a $C_Q$ which can be a good indicator of the image compensation effectiveness. The experiment was conducted through a proposed questionnaire derived from image quality parameters. The experiment was also conducted on paired comparison on Ishihara plates for red-green colour blindness. The results indicate that the method can be adopted to evaluate and rank the effectiveness of colour compensation/correction/simulation algorithms for CVD patients. The proposed method also includes three statistical measures to automatically compare images when colour blind simulation is done with that of a standard tool.

5 | IMPROVEMENTS AND FUTURE WORK

This method can be further improved to include a Z score computation for subjective image quality measurement and SSD method for objective image quality measurement. Further, the sample size for colour blind population can also be increased. The weightage associated with the Ishihara plates and the responses can be optimized using machine learning methods like neural networks and deep learning. Blind quality assessment methods can also be experimented by creating a database of subjective evaluation for colour blindness.

6 | CONCLUSION

There have been several technological developments in the area of digital image processing in computer vision to provide healthcare assistance to colour vision deficient users. The digital solutions primarily involve simulation and colour compensation/correction of colours that CVD patients cannot see. Such compensations should ensure that it does not bring down the quality of the image. Further, the effectiveness of the compensation for the benefit of CVD patients is an important parameter for quality of image in these cases. The proposed study has devised a novel method and a benchmark metric called coefficient of quality ($C_Q$) to measure the image quality in case of compensation/recolouring methods for the CVD patients by
creating a questionnaire based on image quality metrics and parameters that are imperative from a CVD patient’s point of view. This novel method can be an indicator of the quality of colour compensation algorithms for the CVD patients and also provide constructive feedback to authors and researchers to improve the algorithm. The study also looks at a combination of image comparison statistical methods to objectively measure the image quality when there is a reference image present. The experiment and results validate the effectiveness of the method and metric proposed by the authors.

ORCID
Anshu Singla https://orcid.org/0000-0002-1054-8753

REFERENCES
1. Web Content Accessibility Guidelines (WCAG) Overview. https://www.w3.org/WAI/standards-guidelines/wcag/ (2005). Accessed 29 June 2020
2. Color Blindness. https://nei.nih.gov/learn-about-eye-health/eye-conditions-and-diseases/color-blindness (2019). Accessed 08 November 2020
3. Madalena, R., Abel, J.P.G.: Recoloring algorithms for colorblind people: A survey. ACM Comput. Surv. 52(4), 1–37 (2019)
4. Martin, C.E., et al.: Color blindness and a color human visual system model. IEEE Trans. Syst. Man Cybern. 30(4), 494–500 (2000)
5. Ro, Y.M., Yang, S.: Color adaptation for anomalous trichromats. Int. J. Imag. Syst. Techn. 14(1), 16–20 (2004)
6. Brettel, H., et al.: Computerized simulation of color appearance for dichromats. J. Opt. Soc. Am. A. 14(10), 2647–2655 (1997)
7. Stockman, A., et al.: Spectral sensitivities of the human cones. J. Opt. Soc. Am. A. Opt. Imag. Sci. Vis. 10(12), 2491–2521 (1993)
8. Judd, D.B.: The color perceptions of deuteranopic and protanopic observers. J. Opt. Soc. Am. A. 39(3), 252–256 (1949)
9. Chakrabarti, S.: Psychosocial aspects of color vision deficiency: Implications for a career in medicine. Natl. Med. J. India. 31(2), 86–96 (2018)
10. Liu, B., et al.: Efficient image and video re-coloring for colorblindness. Proc. IEEE Intl.Conf. on Multimedia and Expo, New York, NY, (June 2009), pp. 906–909.
11. Machado, G.M., Oliveira, M.M.: Real-time temporal-coherent color contrast enhancement for dichromats. Comput. Graph. Forum. 29(3), 933–942 (2010)
12. Kuhn, G.R., Oliveira, M.M., Fernandes, L.A.F.: An efficient naturalness-preserving image-recoloring method for dichromats. IEEE T Vis. Comput. Gr. 14(6), 1747–1754 (2008)
13. Pinheiro, D.S.W., Lee, J.: An adaptive fuzzy-based system to simulate, quantify and compensate color blindness. ICAIE. 18(1), 29–40 (2011)
14. Utama, D.Q., et al.: RGB color cluster re-coloring algorithm for partial color-blind people. Proc. 5th Intl. Conf. on Instrumentation, Communications, Information Technology, and Biomedical Engineering (ICICI-BME), Bandung, Indonesia (Nov 2018), pp. 219–222
15. Nazrahdin, N.A., Ismal, F.S.: Segmentation and recognition using fuzzy logic human based perception approach for color vision deficiency. Dissertation, UniversitiTeknologi Malaysia, (2015)
16. Iqbal, M., et al.: Image recoloring using spectral decomposition and inpainting. Imaging Sci J 68(3), 188–194 (2020)
17. Chongjin, L., et al.: A modification of CIECAM02 based on the Hunter-Pointer-Estevez matrix. JIST 57(3), 1–8 (2013)
18. Birch, J.: Efficiency of the Ishihara test for identifying red-green color deficiency. Ophthal. Physl. Opt. 17(5), 403–408 (1997)
19. ITU-R BT.500-11: Methodology for the subjective assessment of the quality of television pictures (2002)
20. Mantik, R., et al.: Comparison of four subjective methods for image quality assessment. Comput. Graph. Forum. 31(1), 2478–2491 (2012)
21. ITU-T Recommendation P910: Subjective video quality assessment methods for multimedia applications. (2008)
22. ITU-R Recommendation BT.814-1: Specification and alignment procedures for setting of brightness and contrast of displays. (1994)
23. ITU-R Recommendation BT.1129-2: Subjective assessment of standard definition digital television (SDTV) systems. (1998)
24. Pedram, M., et al.: Subjective and objective quality assessment of image: A survey. J. Electr. Eng. 9(1), 55–83 (2014)
25. Yang, X., et al.: A survey of DNN methods for blind image quality assessment. IEEE Access 7(1), 123788–123806 (2019)
26. Ahhar, S., Wang, Z.: A comprehensive performance evaluation of image quality assessment algorithms. IEEE Access 7(1), 140030–140070 (2019)
27. Jeripothula, A.B., et al.: Blind image quality assessment using a combination of statistical features and CNN. Multimed Tools Appl. 79(1), 23243–23260 (2020)
28. Zhai, G., Min, X.: Perceptual image quality assessment: A survey. Sci. China Inf. Sci. 63(11), 1–52 (2020)
29. He, L., et al.: Objective image quality assessment: A survey. Int. J. Math. 91(11), 2374–2388 (2014)
30. Hou, W., et al.: Blind image quality assessment via deep learning. IEEE Trans. Neural Networks Learn. Syst. 26(6), 1275–1286 (2015)
31. Machado, G.M., et al.: A physiologically-based model for simulation of color 13vision deficiency. IEEE Trans. Vis. Comput. Graph. 15(6), 1291–1298 (2009)
32. Shen, W., et al.: Seamless visual sharing with color vision deficiencies. ACM Trans. Graph. 35(4), 1–12 (2016)
33. Pramoun, T., Amornraksa, T.: Digital watermarking on recolored images for deuteronopia. Proc. 3rd Intl. Conf. on Communication and Information Processing, Tokyo (Nov 2017), pp. 444–448
34. Rasche, K., et al.: Detail preserving reproduction of color images for monochromats and dichromats. IEEE Comput. Graph. Appl. 25(3), 22–30 (2005)
35. Milic, N., et al.: A content-dependent naturalness-preserving daltonization method for dichromatic and anomalous trichromatic color vision deficiencies. J. Imaging. Sci. Techn. 59(1), 1–10 (2015)
36. Ohkubo, T., Kobayashi, K.: A color compensation vision system for colorblind people. Proc. SICE Annual Conf., Chofu, Japan, (Oct 2008), pp. 1286–1289
37. Ribeiro, M.G., Gomes, A.J.P.: A skillet-based recoloring algorithm for dichromats. Proc. IEEE 15th Intl. Conf. on e-Health Networking, Applications and Services (Healthcom 2013), Lisbon, Portugal, (Oct 2013), pp. 702–706
38. Wahab, N., et al.: Color transformation method for protanopia vision deficiency using artificial neural network. JTEC. 8(11), 29–33 (2016)
39. Fikree, M., Paramesran, R.: Naturalness preserving image recoloring method for people with red-green deficiency. Signal Process. Image Commun. 7(1), 126–133 (2016)
40. Larnia, E.A.: Smartphone based image color correction for color blindness. Intl J Mob Tech. 12(3), 104–119 (2018)
41. Yu, M.A., et al.: Color discrimination enhancement for dichromats using self-organizing color transformation. Inf. Sci. 179(6), 830–843 (2009)
42. Huang, J.B., et al.: Information preserving color transformation for protanopia and deuteronopia. IEEE Signal Process. Lett. 14(10), 711–714 (2007)
43. Simon-Liedtke, Joschua, Flatla, David R.: Bakken, Eskild Narum, : et al.: Checklist for daltonization methods: Requirements and characteristics of a good recolouring method. Color Imaging: Displaying, Processing, Hard-copy, and Applications , 21–27 (2017)
44. Vasquez-Espinosa, R.E., Richard, W.C.: The law of comparative judgement: Theory and implementation. Louisiana state university Baton Rouge remote sensing and image processing lab. (October 1982)
45. Kotz, S., et al.: Encyclopedia of Statistical Sciences. 16, 2 John Wiley & Sons, Hoboken (2006)
46. Watters, Paul Andrew, Boslaugh, Sarah: Statistics in a Nutshell, 2nd ed., 1, 1 O’Reilly, Newton (2008)
47. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vis. 59(2), 91–110 (2004)
How to cite this article: S M, Singla A. A novel image quality assessment method and coefficient of quality for digital solutions of colour blindness. *IET Image Processing*. 2021;1–13. https://doi.org/10.1049/ipr2.12213.

APPENDIX

| Variable name/acronym | Description/full form |
|-----------------------|-----------------------|
| $Q_w$                 | Weighted mean score for each question |
| $J_w$                 | Weightage given by judges for a question |
| $m$                   | Number of judges |
| $R_w$                 | Weighted mean response weightage score |
| $I_w$                 | Weighted mean ideal weightage |
| $P_i$                 | Weightage given to the image |
| $S_r$                 | Score values against the responses provided by the evaluators for reference image |
| $S'_r$                | Score values against the responses provided by the evaluators for colour compensated image |
| $R_w$                 | Weighted mean response weightage score calculated for the reference image |
| $R'_w$                | Weighted mean response weightage score calculated for the colour compensated image |
| $CQ$                  | Coefficient of quality |
| MSR                   | Mean significant ratio |
| SSD                   | Sum of the squares of the deviations |
| BIQA                  | Blind image quality assessment |
| NIQA                  | Natural image quality assessment |
| ITU                   | International Telecom Union |
| CVD                   | Colour vision deficiency |
| IQA                   | Image quality assessment |
| SIFT                   | Scale invariant feature transform |