Article

Reduced Reference Quality Assessment for Image Retargeting by Earth Mover's Distance

Longsheng Wei ^{1,2,3} , Lei Zhao ^{1,3} and Jian Peng ^{1,3,*}

¹ School of Automation, China University of Geosciences, Wuhan 430074, China; weilongsheng@cug.edu.cn (L.W.); dttutty@gmail.com (L.Z.)

² Key Laboratory of Geological Survey and Evaluation of Ministry of Education, Wuhan 430074, China

³ Hubei Key Laboratory of Advanced Control and Intelligent Automation for Complex Systems, Wuhan 430074, China

* Correspondence: pengjian@cug.edu.cn

Abstract: A reduced reference quality assessment algorithm for image retargeting by earth mover's distance is proposed in this paper. In the reference image, all the feature points are extracted using scale invariant feature transform. Let the histograms of image patch around each feature point be local information, and the histograms of saliency feature as global information. Those feature information is extracted at the sender side and transmitted to the receiver side. After that, the same feature information extraction process is performed for the retargeted image at the receiver side. Finally, all feature information of the reference and retargeted images is used collectively to compute the quality of the retargeted image. An overall quality score is calculated from the local and global similarity measure using earth mover's distance between reference and retargeted images. The key step in our algorithm is to provide an earth mover's distance metric in a manner that indicates how the local and global information in the reference image is preserved in corresponding retargeted image. Experimental results show that the proposed algorithm can improve the image quality scores on four common criteria in the retargeted image quality assessment community.

Keywords: reduced reference; image quality assessment; earth mover's distance; human visual perception

Citation: Wei, L.; Zhao, L.; Peng, J. Reduced Reference Quality Assessment for Image Retargeting by Earth Mover's Distance. *Appl. Sci.* **2021**, *11*, 9776. <https://doi.org/10.3390/app11209776>

Academic Editor: Antonio Fernández

Received: 8 September 2021

Accepted: 16 October 2021

Published: 19 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Image quality is a basic concept in many image processing and computer vision applications, such as acquisition, transmission, and display. With advances in information technology and visual communication, assessing image quality has become a fundamental and challenging problem. Image Quality Assessment (IQA) automatically measures the image visual quality by effective computational models [1]. The image quality assessment (IQA) approach attempts to estimate the image quality based on human visual perception in an objective manner. Most IQA models based on full reference (FR) and achieved very good results, while most no reference (NR) IQA methods are designed for some predefined specific distortion types. Reduced reference (RR) IQA algorithms provide a proper compromise between FR and NR approaches, and they estimate the image quality with limited access to the reference image [2].

Recently, multimedia retargeting has attracted much attention in a graph and vision research. Retargeting techniques for image and video utilize the original visual scene to display that scene at different sizes or aspect ratios on different display screens. Many retargeting models have been proposed, such as multi-operator (MO), cropping (CR), streaming video (SV), shift-map (SM), seam carving (SC), scaling (SCL), scale-and-stretch (SNS), and warping (WARP) [3]. These models are either pixel-based or patch-based. Nevertheless, the FR image is not often available in time of retargeted IQA, but the reference image may be briefly described through partial information. Thus, an RR quality assessment criterion

is needed to employ. Therefore, how to effectually assess the quality of the RR retargeted images is an important and challenging problem.

However, general IQA approaches, such as peak signal-to-noise ratio (PSNR) and structural similarity (SSIM), are not applied to the retargeted images because those algorithms require the coupled images have the same sizes. The existing retargeted IQA algorithms mainly based on full reference, such as color layout descriptor (CL), edge histogram (EH), bidirectional warping (BDW), bidirectional similarity (BDS), and scale invariant feature transform (SIFT) Flow [3]. Those algorithms have been used to evaluate retargeted images quality, although their results are not always consistent with subjective image evaluations. Since humans are the final evaluators of image quality, the goal must be to develop an automatic method that can evaluate image quality in a way that will be in agreement with subjective evaluation. RR algorithms often assess image quality by comparing the same features between reference and retargeted images, such as discrete cosine transform (DCT) [4], visual information fidelity (VIF) [5], divisive normalization transformation (DNT) [6], reduced reference entropic difference (RRED) [7], and reduced-reference SSIM (RRSSIM) [8]. Bampis et al. [9] proposed Gaussian scale mixture models to assess a RR image by computing the weighted entropies between reference and retargeted images. According to different types and levels of degradation, which could strongly influence saliency detection, a saliency-induced RR IQA method was introduced by Min et al. [10]. However, when the image is retargeted, some features may also change, so it is hard to compare our proposed algorithm with them directly. Thus, this paper proposes a useful method to estimate the visual quality for RR retargeted images.

Since the size of the reference image is inconsistent with that of the retargeted image, this paper extracts the corresponding feature points of the two images and calculates the EMD between the image blocks corresponding to the feature points as local information. Meanwhile, this paper extracts the visual saliency features of the two images and calculates the EMD between the corresponding features as global information. Finally, local and global information are combined to generate a final quality score. Our proposed algorithm can obtain high quality prediction accuracy with the limited amount of RR features. This algorithm is illustrated in Figure 1.

Figure 1. The framework of RR retargeted IQA by EMD.

2. EMD between Two Histograms

When the sizes of the reference image and the retargeted image are different, we cannot compare the two images directly. However, the histogram of image is independent of the size of the image, so we can use EMD to compute the similarity between reference image and retargeted image. We introduce the EMD metric between two given normalized histograms with the same number of bins in this part.

2.1. Basic EMD

Rubner et al. [11] first introduced EMD for measuring the difference between texture and color, in which the EMD was applied to the distributed signatures instead of histograms directly. As a matter of fact, a histogram can also be considered to be a special kind of signatures.

The classical EMD between two histograms is the lowest cost of transporting one histogram into another, in which the cost is usually defined as that the amount of weight multiplies the ground distance between two histograms. This formalization is easily generalized to two normalized histograms with the same number of bins [12].

Given two n -bin histograms, $H^1 = \{h_i^1, i = 1, 2, 3, \dots, n\}$ and $H^2 = \{h_j^2, j = 1, 2, 3, \dots, n\}$, H^1 is transformed into H^2 by moving “mass” from h_i^1 to h_j^2 for every pair of (i, j) , such that the difference of two histograms is minimized. Let another n -bin all-zero histogram be T , and we denote the flow f_{ij} as the amount, which is moved from the i -th bin in H^1 to the j -th bin in T . Then, we can define the EMD metric between H^1 and H^2 to be the minimum amount flow that is demanded to make the histogram of T to be identical with H^2 . Therefore, the EMD between H^1 and H^2 is expressed in mathematically as follows:

$$\text{EMD}(H^1, H^2) = \min_{\{f_{i,j}, i,j=1,2,3,\dots,n\}} \sum_{i=1}^n \sum_{j=1}^n f_{i,j} d_{i,j}, \quad (1)$$

subject to $\sum_{j=1}^n f_{i,j} = h_i^1$, $\sum_{i=1}^n f_{i,j} = h_j^2$, $f_{i,j} \geq 0$ and $i, j = 1, 2, 3, \dots, n$, where $d_{i,j}$ is the ground distance between bin i and bin j . Let $d_{i,j} = |i - j|$ to be the L^1 distance in this section.

2.2. Weighted EMD

For each given reference and retargeted images, we obtain all the matching point pairs. We take an image patch around each matching point in the point pairs.

If we use the above EMD as the distance measure in two histograms of image patch directly, we do not consider the spatial information of pixels in the patch. In fact, pixels near the center of matching point are more important than the others for computing similarity [13]. To solve this problem, we bring the weight to the pixels of different locations, and then we can calculate the EMD between the weighted histograms.

Because the importance of the pixel is inversely proportional to the distance between the pixel and center point, we define the normalized weight value $w(\cdot)$ for every pixel in the image patch by

$$w(i) = \begin{cases} \frac{1}{d(i)} / \sum_{j \in S} \frac{1}{d(j)} & \text{otherwise} \\ 1 & i \text{ is the center} \end{cases}, \quad (2)$$

where $d(i)$ is the ground distance from the pixel i to center matching point in the patch S . By using the weight value as the pixel's contribution to their histogram bins, we construct weighted histograms, and, by applying those weighted histograms to EMD, we can more accurately calculate the distance between the two image patches.

In order to speed up the calculation, we choose the L^1 -based distance as the ground distance $d_{i,j}$. With this choice, Levina has proven that the EMD between normalized weighted histograms equals to linear Wasserstein distance [14]. Under those conditions, the EMD can be written as:

$$\text{EMD}(IH^1, IH^2) = \sum_{i=1}^n \left| \sum_{j=1}^i IH^1(j) - \sum_{j=1}^i IH^2(j) \right|, \quad (3)$$

where IH^1 and IH^2 are two normalized weighted histograms.

3. Quality Assessment Using EMD

In this section, the SIFT and saliency features are extracted from reference and retargeted images, respectively. Then, the local and global similarity measures are computed by those features. Last, the overall quality score is obtained by fusing the local and global EMD.

3.1. Local EMD Based on SIFT Features

People usually hope that retargeted images can accurately preserve the local structural information in corresponding regions in the reference images. If we establish the pixel correspondence, we can compare the structural information in corresponding local regions between two images. Thus, implementation of this pixel correspondence between images is a key factor in quality assessment.

Therefore, we should build a matching algorithm between pixels in the reference image and retargeted image. The SIFT descriptor is widely used in pixel matching between different size scenes [15], and it has been verified to be very effective in matching areas of corresponding dense in images [16]. As in optical flow algorithms, SIFT flow utilizes SIFT descriptors, rather than original pixels, to match image densities.

Firstly, in the reference image, we extract all SIFT feature points and their eigenvalues; then, we get an image patch around each feature point, and transform the image patch into a histogram. Secondly, we transfer all the eigenvalues and corresponding image patch histograms through an ancillary channel. Thirdly, in the retargeted image, we also extract all SIFT feature points and their eigenvalues; then, we find all matching points by the feature matching, and we can easily get image patch histograms around each matching point. Lastly, we use above EMD to calculate these two histograms centered on feature points. In fact, we do not match these points directly; rather, we just find the corresponding SIFT feature points in the retargeted image from reference image.

The local image quality can be expressed by all the matching image patches, so we calculate the local image quality score using EMD (LEMD) by average strategy:

$$\text{LEMD} = \frac{1}{M} \sum_{i=1}^M \text{EMD}(IH_{p,i}^{\text{ref}}, IH_{p,i}^{\text{ret}}), \quad (4)$$

where M is the total number of matching points, and $\text{EMD}(IH_{p,i}^{\text{ref}}$ and $IH_{p,i}^{\text{ret}})$ are the i -th matching image patch normalized weighted histogram for reference and retargeted images, respectively.

3.2. Global EMD Based on Saliency Features

Although the image patches are well-evaluated for local quality, the quality of the image includes not only local quality but also global quality. Previous studies show that people can use the eye tracking data to promote the effect of IQA metrics [17], while humans are the ultimate evaluators of image quality, so it is reasonable to introduce principles of human visual saliency features into image evaluation as the global measurement. This can make objective image evaluation more consistent with subjective image evaluation.

In Itti's model [18], color, intensity, and orientation are extracted for visual saliency features. Since human vision is sensitive to image texture information, texture feature is added in this section. In this section, we extract ten visual saliency features from reference image, including two color contrasts (red-green and blue-yellow), two intensity contrasts (light to dark and dark to light), four orientation features ($0^\circ, 45^\circ, 90^\circ, 135^\circ$), and two texture features (original and extended LBP) [19].

In the reference image, we extract ten saliency features, and we turn each feature map into a histogram. We transfer all the histograms. In the retargeted image, we also extract ten saliency features and turn each feature map into a histogram. Since the center of the image is the center of human vision, we make the center of each feature map as the center

of the histogram. Therefore, we obtain global image quality score using EMD (GEMD) by average strategy:

$$\text{GEMD} = \frac{1}{N} \sum_{j=1}^N \text{EMD}(IH_{f,j}^{\text{ref}}, IH_{f,j}^{\text{ret}}), \quad (5)$$

where N is the total number of saliency features, and $IH_{f,j}^{\text{ref}}$ and $IH_{f,j}^{\text{ret}}$ are the j -th saliency features normalized weighted histogram for reference and retargeted images, respectively.

3.3. The Overall Quality Score Based on Local and Global EMD

The LEMD proposes a local metric about how much of the structure information is preserved in corresponding image patches, and the GEMD suggests a global method about how similarity of visual saliency features between reference and retargeted images. Therefore, we define an overall quality score (QS) based on local and global EMD as follows:

$$\text{QS} = \text{LEMD} + \text{GEMD} = \frac{1}{M} \sum_{i=1}^M \text{EMD}(IH_{p,i}^{\text{ref}}, IH_{p,i}^{\text{ret}}) + \frac{1}{N} \sum_{j=1}^N \text{EMD}(IH_{f,j}^{\text{ref}}, IH_{f,j}^{\text{ret}}). \quad (6)$$

According to the above EMD calculation process, we can find that the value of EMD is the distance between two image patches. The smaller value of QS, the closer the distance, and the more similarities between the two images.

4. Experimental Results

In the field of IQA, indeed, there are many public databases, such as TID2013 [20] and KADID-10K [21], CID2013 [22], LIVE challenge [23], and KonIQ-10K [24]. However, these databases are either synthetically or authentically distorted IQA databases. We aim at assessing the retargeted images database in this paper, so we choose a popular public database [3] for retargeted images to validate the proposed algorithm. The database contains 37 source reference images and corresponding retargeted images. Two hundred and ten participants take part in the assessment of subjective image quality. All the data constitute this database. Every source reference image has been retargeted by eight different models, including MO, CR, SV, SM, SC, SCL, SNS, and WARP. Furthermore, subjective image evaluation results, for all retargeted images using above models, are also provided in this database. Figure 2 shows an example of child image about a source reference image and eight retargeted images with different models.

Figure 3 shows the histogram of subjective and objective assessment. The left is the subjective votes for the child image in Figure 2 by participants, and the right is objective results by our proposed algorithm on the same image. For the image obtained by each retargeted model, every participant gives a score by observing the image itself, and the score range is [0, 100]. Then, the subjective score is the mean value of all participants' scores. The objective score is calculated by our proposed algorithm in Equation (6). There is no specific objective score range, and the score is the sum of local EMD and global EMD. Therefore, both histograms are in different scales. Because our algorithm is based on EMD, which is the distance between two image patches, the smaller value of objective result, the closer the distance, and the more similarities between the two images. When calculating the correlation, we only calculate the consistency of subjective and objective scores, and we do not evaluate the values of the two scores.

The existing retargeted IQA algorithms mainly based on full reference, while existing RR IQA algorithms are often used for some special distorted images, such as noise, blur, and JPG, not for retargeted images, so it is hard to compare our proposed algorithm with them directly. Therefore, we separate all the compared IQA algorithms into two types in the experiment. One is full reference IQA algorithms for retargeted images, including CL, EH, BDW, BDS, and SIFT Flow; the other is RR IQA algorithms for non-retargeted images, including MA, VIF, DNT, RRED, and RRSSIM.

Figure 2. An example of a source reference image and eight retargeted images with different models.

Figure 3. Histogram of subjective and objective assessment. (left) Subjective votes for the child image by participants. (right) Objective results by our proposed algorithm on the same image.

Four criteria [8] are used to assess how well the objective quality scores can predict the subjective quality scores: (1) Pearson linear correlation coefficient (PLCC), (2) Spearman's rank correlation coefficient (SRCC), (3) Kendall's rank correlation coefficient (KRCC), and (4) Root mean squared (RMS) error.

In the experiment, for each reference image and a series of retargeted images, the subjective rankings can be obtained from the database, and the different objective rankings can be computed by different IQA algorithms, so those algorithms can be compared by the subjective and objective correlation.

To better analyze the effectiveness of all the algorithms for different types of image assessment, 37 images in the database are divided into 6 types according to the selected attributes: 25 lines or edges images, 15 faces or people images, 6 texture images, 18 foreground objects images, 16 geometric structures images, and 6 symmetry images. Note that one image may belong to several different types, since it can contain several attributes, such as faces and people, which often belong to the foreground objects images.

Table 1 presents these correlation scores of subjective and objective measures for the KRCC according to image attributes. As expected, the FR IQA for retargeted images

algorithms show smaller scores correspondence with the subjective assessment, although SIFT Flow and EH achieves higher scores for images classified as containing apparent texture and geometric structures. RR IQA for non-retargeted images algorithms show some better than FR IQA for retargeted images. The near-zero correlation for nearly all image types suggests they cannot predict well the subjective assessment. Their unsatisfying performance has to do with both the image features they use for measuring the distance, and with the way they construct correspondence between the images.

Table 1. Correlation scores of subjective and objective measures for the KRCC according to image attributes. Highest score in each column appears in bold.

	Algorithm	Lines/Edges	Faces/People	Texture	Foreground Objects	Geometric Structures	Symmetry
FR IQA for retargeted images	EL	0.043	−0.076	−0.060	−0.079	0.103	0.298
	CL	−0.023	−0.181	−0.071	−0.183	−0.009	0.214
	BDW	0.031	0.048	−0.048	0.060	0.004	0.119
	BDS	0.040	0.190	0.060	0.167	−0.004	−0.012
	SIFT Flow	0.097	0.252	0.119	0.218	0.085	0.071
RR IQA for non-retargeted images	DCT	0.124	−0.002	0.051	0.201	0.035	0.103
	VIF	0.210	0.233	0.078	0.230	0.101	−0.004
	DNT	0.142	0.248	−0.025	0.185	0.046	0.151
	RRED	−0.031	0.341	0.075	0.284	−0.003	0.293
	RRSSIM	0.179	0.328	0.020	0.257	0.237	0.226
RR IQA for retargeted images	LEMD	0.236	0.432	0.068	0.379	0.026	0.332
	GEMD	0.125	0.452	0.025	0.420	0.012	0.150
	QS	0.262	0.461	0.073	0.438	0.034	0.357

The performance comparison with different IQA algorithms is given in Table 2. It is easy to find that our method is better than the given FR IQA for retargeted images and the given RR IQA for non-retargeted images in Table 2.

PLCC is linear correlation metric, SRCC and KRCC are nonparametric rank correlation metric, and the agreement range is $[-1, 1]$. In this way, higher correlation coefficient predicates higher sequences agreement. RMS is an error metric, so lower error value represents higher agreement.

By comparing the image patches around SIFT feature points between reference and retargeted images, EMD provided a more robust metric of their similarity. As a result, our algorithm was able to obtain high quality prediction accuracy with the limited amount of RR features, and our algorithm could achieve good results in experiment.

Table 2. The performance comparison with different IQA algorithms.

	Algorithm	PLCC	SRCC	KRCC	RMS
FR IQA for retargeted images	EL	0.048	0.025	0.004	0.324
	CL	0.019	0.013	−0.068	0.457
	BDW	0.062	0.050	0.046	0.274
	BDS	0.091	0.097	0.083	0.262
	SIFT Flow	0.149	0.151	0.145	0.238
RR IQA for non-retargeted images	DCT	0.073	0.069	0.056	0.266
	VIF	0.185	0.187	0.168	0.232
	DNT	0.161	0.153	0.137	0.241
	RRED	0.253	0.247	0.234	0.225
	RRSSIM	0.260	0.266	0.259	0.221
RR IQA for retargeted images	LEMD	0.325	0.322	0.294	0.172
	GEMD	0.129	0.126	0.120	0.236
	QS	0.370	0.365	0.331	0.164

5. Discussion

We can find the overall result of correlation is between 0.331 and 0.370 in Table 2. There are three main reasons why the correlation is so low.

(1) The retargeted image database is complex. This database includes lots of irregularly-textured areas, such as grass, water, or trees, and the images in database are retargeted by removing, inserting, or optimizing pixels (or patches) to preserve image content, so the size of the retargeted image is very different from the original image. By this way, the images contains either dense information or global and local structures that may be damaged during resizing.

(2) PLCC, SRCC, and KRCC are three main statistical correlation coefficients, which describe the linear or rank correlation between subjective score and algorithm score. However, the subjective score of retargeted images is obtained by averaging many scores according to 210 participants, which affects the linearity or rank of subjective score and algorithm score.

(3) This paper uses a reduced reference retargeted image quality assessment algorithm, which only uses part of the reference image information, rather than all the information, so the correlation is relatively low.

We could find that our algorithm was better than all the given FR IQA algorithms for retargeted images. The reason was that our algorithm not only considered the local metric (Local EMD) but also considered the global saliency features (Global EMD), while the FR IQA algorithms for retargeted images only used the local metrics. We also found that our algorithm was better than all the given RR IQA for non-retargeted images algorithms. That was because our algorithm did not depend on the same size between reference image and retargeted image, while the compared RR IQA algorithms often need the same sizes between two images. When the two sizes were different, the non-retargeted algorithms would compare their common regions directly. As a matter of fact, they only evaluated a part of images.

Because GEMD just used ten visual saliency features, its evaluation results were not better than many other algorithms. However, it provided a good complement to LEMD and improved the overall evaluation results.

6. Conclusions

In this paper, we have proposed an RR retargeted IQA algorithm using EMD. Each reference image is retargeted through a retargeting channel, and the local and global information, which usually has fewer data than reference image, is transferred through a specific ancillary channel. We extracted SIFT, image block histograms, and saliency feature histograms for corresponding reference and retargeted images, respectively. The overall quality score is calculated from those feature information using EMD between reference and retargeted images. Experimental results demonstrated that the comparison indexes of the proposed algorithm were better than the indexes of given algorithms.

The key step in our algorithm is to provide an EMD metric in a manner that indicates how the local and global information in the reference image is preserved in the corresponding retargeted image. In future works, multi-scale EMD method will be added to extend the RR retargeted IQA approach.

Author Contributions: Conceptualization, L.W. and J.P.; methodology, L.W.; software, L.Z.; validation, L.Z. and L.W.; formal analysis, L.Z.; investigation, L.W.; resources, J.P.; data curation, L.W. and J.P.; writing—original draft preparation, L.W.; writing—review and editing, L.Z.; visualization, L.Z.; supervision, J.P.; project administration, L.W. and J.P.; funding acquisition, L.W. and J.P. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Opening Fund of Key Laboratory of Geological Survey and Evaluation of Ministry of Education (GLAB2020 ZR06) and the Fundamental Research Funds for the Central Universities, the Joint Foundation of China Aerospace Science and Industry for

Equipment Pre-Research 2020, and the National Natural Science Foundation of China under contracts on 61603357.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

The following abbreviations are used in this manuscript:

Abbreviation	Full Name
RR	Reduced Reference
EMD	Earth Mover's Distance
SIFT	Scale Invariant Feature Transform
IQA	Image Quality Assessment
FR	Full Reference
NR	No Reference
MO	Multi-operator
CR	Cropping
SV	Streaming Video
SM	Shift-map
SC	Seam Carving
SCL	Scaling
SNS	Scale-and-stretch
WARP	Warping
PSNR	Peak Signal-to-noise Ratio
SSIM	Structural Similarity
CL	Layout Descriptor
EH	Edge Histogram
BDW	Bidirectional Warping
BDS	Bidirectional Similarity
VIF	Information Fidelity
DNT	Divisive Normalization Transformation
RRED	Reduced Reference Entropic Differencing
RRSSIM	Reduced-Reference SSIM
LEMD	Local Image Quality Score using EMD
GEMD	Global Image Quality Score using EMD
QS	Quality Score
PLCC	Pearson Linear Correlation Coefficient
SRCC	Spearman's Rank Correlation Coefficient
KRCC	Kendall's Rank Correlation Coefficient
RMS	Root Mean Squared

References

1. Zhou, Z.; Li, J.; Quan, Y.; Xu, R. Image Quality Assessment Using Kernel Sparse Coding. *IEEE Trans. Multimed.* **2021**, *23*, 1592–1604. [\[CrossRef\]](#)
2. Liu, Y.; Zhai, G.; Gu, K.; Liu, X.; Zhao, D.; Gao, W. Reduced-Reference Image Quality Assessment in Free-Energy Principle and Sparse Representation. *IEEE Trans. Multimed.* **2018**, *20*, 379–391. [\[CrossRef\]](#)
3. Rubinstein, M.; Gutiérrez, D.; Sorkine, O.; Shamir, A. A Comparative Study of Image Retargeting. *ACM Trans. Graph.* **2010**, *29*, 1–9. [\[CrossRef\]](#)
4. Ma, L.; Li, S.; Zhang, F.; Ngan, K.N. Reduced-Reference Image Quality Assessment Using Reorganized DCT-Based Image Representation. *IEEE Trans. Multimed.* **2011**, *13*, 824–829. [\[CrossRef\]](#)
5. Wu, J.; Lin, W.; Shi, G.; Liu, A. Reduced-Reference Image Quality Assessment With Visual Information Fidelity. *IEEE Trans. Multimed.* **2013**, *15*, 1700–1705. [\[CrossRef\]](#)
6. Li, Q.; Wang, Z. Reduced-Reference Image Quality Assessment Using Divisive Normalization-Based Image Representation. *IEEE J. Sel. Top. Signal Process.* **2009**, *3*, 202–211. [\[CrossRef\]](#)

7. Soundararajan, R.; Bovik, A.C. RRED Indices: Reduced Reference Entropic Differencing for Image Quality Assessment. *IEEE Trans. Image Process.* **2012**, *21*, 517–526. [[CrossRef](#)] [[PubMed](#)]
8. Rehman, A.; Wang, Z. Reduced-Reference Image Quality Assessment by Structural Similarity Estimation. *IEEE Trans. Image Process.* **2012**, *21*, 3378–3389. [[CrossRef](#)] [[PubMed](#)]
9. Bampis, C.G.; Gupta, P.; Soundararajan, R.; Bovik, A.C. SpEED-QA: Spatial Efficient Entropic Differencing for Image and Video Quality. *IEEE Signal Process. Lett.* **2017**, *24*, 1333–1337. [[CrossRef](#)]
10. Min, X.; Gu, K.; Zhai, G.; Hu, M.; Yang, X. Saliency-induced reduced-reference quality index for natural scene and screen content images. *Signal Process.* **2018**, *145*, 127–136. [[CrossRef](#)]
11. Rubner, Y.; Tomasi, C.; Guibas, L.J. The Earth Mover's Distance as a Metric for Image Retrieval. *Int. J. Comput. Vis.* **2000**, *40*, 99–121. [[CrossRef](#)]
12. Ling, H.; Okada, K. An Efficient Earth Mover's Distance Algorithm for Robust Histogram Comparison. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *29*, 840–853. [[CrossRef](#)] [[PubMed](#)]
13. Lin, Y.; Tang, Y.Y.; Fang, B.; Shang, Z.; Huang, Y.; Wang, S. A Visual-Attention Model Using Earth Mover's Distance-Based Saliency Measurement and Nonlinear Feature Combination. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 314–328. [[CrossRef](#)] [[PubMed](#)]
14. Levina, E.; Bickel, P. The Earth Mover's distance is the Mallows distance: Some insights from statistics. In Proceedings of the Eighth IEEE International Conference on Computer Vision (ICCV), Vancouver, BC, Canada, 7–14 July 2001; Volume 2, pp. 251–256. [[CrossRef](#)]
15. Ma, J.; Qiu, W.; Zhao, J.; Ma, Y.; Yuille, A.L.; Tu, Z. Robust L2E Estimation of Transformation for Non-Rigid Registration. *IEEE Trans. Signal Process.* **2015**, *63*, 1115–1129. [[CrossRef](#)]
16. Liu, C.; Yuen, J.; Torralba, A. SIFT Flow: Dense Correspondence across Scenes and Its Applications. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 978–994. [[CrossRef](#)] [[PubMed](#)]
17. Xue, W.; Zhang, L.; Mou, X.; Bovik, A.C. Gradient Magnitude Similarity Deviation: A Highly Efficient Perceptual Image Quality Index. *IEEE Trans. Image Process.* **2014**, *23*, 684–695. [[CrossRef](#)] [[PubMed](#)]
18. Baluch, F.; Itti, L. Mining videos for features that drive attention. In *Multimedia Data Mining and Analytics*; Springer: Berlin, Germany, 2015; pp. 311–326.
19. Wei, L.; Luo, D. A biologically inspired computational approach to model top-down and bottom-up visual attention. *Opt. Int. J. Light Electron. Opt.* **2015**, *126*, 522–529. [[CrossRef](#)]
20. Ponomarenko, N.; Jin, L.; Ieremeiev, O.; Lukin, V.; Egiazarian, K.; Astola, J.; Vozel, B.; Chehdi, K.; Carli, M.; Battisti, F. Image database TID2013: Peculiarities, results and perspectives. In Proceedings of the Signal Processing: Image Communication, Bellingham, WA, USA, 31 July–4 August 2005; Volume 32, pp. 57–77. [[CrossRef](#)]
21. Lin, H.; Hosu, V.; Saupe, D. KADID-10k: A Large-scale Artificially Distorted IQA Database. In Proceedings of the 2019 Eleventh International Conference on Quality of Multimedia Experience (QoMEX), Berlin, Germany, 5–7 June 2019; pp. 1–3. [[CrossRef](#)]
22. Virtanen, T.; Nuutinen, M.; Vaahteranoksa, M.; Oittinen, P.; Häkkinen, J. CID2013: A database for evaluating noreference image quality assessment algorithms. *IEEE Trans. Image Process.* **2015**, *24*, 390–402. [[CrossRef](#)] [[PubMed](#)]
23. Ghadiyaram, D.; Bovik, A.C. Massive Online Crowdsourced Study of Subjective and Objective Picture Quality. *IEEE Trans. Image Process.* **2016**, *25*, 372–387. [[CrossRef](#)] [[PubMed](#)]
24. Lin, H.; Hosu, V.; Saupe, D. Koniq-10k: Towards an ecologically valid and large-scale iqa database, Pittsburgh, Pennsylvania. *arXiv* **2018**, arXiv:1803.08489.