ISSN (Print) : 0974-6846
ISSN (Online) : 0974-5645
Indian Journal of Science and Technology, *Vol 8(S7), 117–121, April 2015*

Performance Evaluation of Image Retrieval System Based on Error Metrics

Vijay Jeyakumar* and Bommanna Raja Kanagaraj

Centre for System Design and Information Processing, Department of Biomedical Engineering, PSNA College of Engineering and Technology, Dindigul, India; ece.vijay@gmail.com, dr.k.bommannaraja@gmail.com

Abstract

Content based image classification/retrieval system has become a widely useful component for any kind of database systems. However, the existing methods do not provide adequate information to the user query based on feature descriptors. In particular, retrieval in medical environment require special attention in feature extraction and feature selection. As, there is no benchmark retrieval algorithm proposed to a database system which contains all kind of medical images, time has come to evaluate the existing system methodology and to devise an interactive and effective retrieval system. Hence, various mathematical and human visual system (HVS) based measurement methods are applied to evaluate the relevancy between the query and retrieved images. From the obtained results and comparisons, the feature based retrieval system does not exploit in a better way to meet the requirements of medical users.

Keywords: CBIR, DICOM, Error Metrics, Retrieval, Structural Similarity

1. Introduction

Due to steadily growing various multimedia data such as video, audio, medical and commercial data, many algorithms and tools are proposed by various researchers to find the exact match for the given query. But still, no general architecture has been formulated to handle different databases with varying characteristics. Finding the suitable solution for this particular problem is still open and eagerly waiting for benchmark solutions to fill the void[1]. Over the past four decades, Content Based Visual Information Retrieval (CBVIR) or Content Based Image Retrieval (CBIR) plays a major role in the field of computer vision. Now a day, in hospitals digital images are produced in ever increasing quantities which is used for diagnosis and therapy. In many articles, content based access to medical images for automated clinical decision making are proposed[2]. Only very few systems are accepted by the medical practitioners for their applications and rest of the systems are totally isolated from clinical practice. A common question arises among all the readers are whether CBIR belongs to an information retrieval task or a classification task. As many similarities are seen among them, the major principle difference is, Classification tasks have a limited number of classes of topics/items and training data for each of the classes that allow training of class-specific parameters; Retrieval tasks have no fixed classes of items/objects in the database and usually no training data available; documents can be relevant for a particular retrieval task or information need, with relevance being potentially user-dependent. The classification paradigm follows the machine learning techniques to find the task, whereas the retrieval paradigm follows the general information retrieval. At the end, when used for CBIR, both of the paradigms produce the images by visually similar features. As the resultant images are viewed by human, the evaluations of visually similar images are done by subjective evaluation. As *subjective* evaluation is too time consuming process and inconvenient, *objective* image quality assessment method is required. Objective image quality metric can be used to adjust image quality, to optimize algorithms and to benchmark image processing systems and algorithms[3]. Objective image quality metrics can be classified as follows: When a complete reference image is compared with the distorted images is known as *Full – Reference* and when the reference is not available to evaluate is termed as *No – Reference*. If the reference image is available partially

**Author for correspondence*

with its extracted feature values to evaluate the distorted image is known as *Partially – Reference*. Apart from this, two classes of objective quality measurement methods are derived. They are mathematically defined measures such as Mean Squared Error (MSE), Signal to Noise Ratio (SNR), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE) and its derivatives. The second class of measurement method focuses on HVS characteristics[4,5]. Though mathematical measures are used widely because of its low computational complexity, many HVS based measures are also proposed to view images when in different viewing conditions. Hence in image retrieval applications, neither mathematical measures nor HSV measures alone not sufficient to evaluate. In this paper various error metric measures are considered and its characteristics are compared in medical image retrieval system environment.

2. Methodology

The most widely used *full - reference* image quality assessment method to evaluate in the easier way is MSE, computed by averaging the squared intensity differences of query image and reference image pixels.

$$\mathrm{MSE} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{i=1}^{N}(x(i,j) - y(i,j))^2 \tag{1}$$

Where x(i, j) is the original image (reference) and y(i, j) is modified image (distorted). M and N are the number of pixels of reference and distorted images respectively.

Similarly Mean Absolute Error (MAE) can be calculated as follows

$$\mathrm{MAE} = \frac{1}{MN}\sum_{i=1}^{M}\sum_{i=1}^{N}|x(i,j) - y(i,j)|. \tag{2}$$

From the above equations 1 & 2, Root mean square error (RMSE), root mean absolute percentage error (RMAPE), root mean square percentage error (RMSPE) can be estimated.

MSE is usually converted into peak signal to noise ratio (PSNR) as

$$\mathrm{PSNR} = 10\log_{10}\frac{L^2}{\mathrm{MSE}} \tag{3}$$

Where, L is the dynamic range of image pixel intensities.

Though the above equations are convenient mathematically, simple, convexity, symmetry and differentiability, they are not used to perceived visual quality[6-8].

Visual signal to noise ratio (VSNR) is used to estimate the visual fidelity of images based on their near-supra threshold properties of subjective vision. The computation is based on wavelet based on physical luminance and angle to incorporate multiple viewing conditions[9].

Weighted peak signal to noise ratio is described as follows

$$\mathrm{WPSNR} = 10\log_{10}\frac{(255)^2}{\mathrm{WMSE}} \tag{4}$$

$$\mathrm{WMSE} = \frac{\mathrm{MSE}}{[1+\mathrm{var}(i,j)]^2} \tag{5}$$

Where, var(i, j) is the maximum local variance of an image.

SSIM[10] algorithm contains three terms which analyses the similarity between two images in the aspects of luminance $l(x, y)$, contrast c(x, y) and structure s(x, y)

$$\begin{aligned}\mathrm{SSIM}(x,y) &= l(x,y)\cdot c(x,y)\cdot s(x,y)\\ &= \left(\frac{2\mu_x\mu_y + C_1}{\mu_2^x + \mu_2^y + C_1}\right)\cdot\left(\frac{2\sigma_x\sigma_y + C_2}{\sigma_2^x + \sigma_2^y + C_2}\right)\cdot\left(\frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}\right)\end{aligned} \tag{6}$$

Where, $C_1 = (K_1L)^2$, $C_2 = (K_2L)^2$ and $C_3 = C_2/2$; L is the dynamic range of pixel intensity values and $K_1 << 1$ and $K_2 << 1$ are scalar constants.

As the image quality perceived by the observer is varied due to multi-scale attributes, distortion and change of distance, Multi-scale SSIM (MS-SSIM)[11] is performed to assess the image quality on multiple scales.

$$\mathrm{MS-SSIM}(x,y) = [l_M(x,y)^{\alpha_M}\prod_{j=1}^{M}[c_j(x,y)]^{\beta_j}[s_j(x,y)]^{\gamma_j} \tag{7}$$

At each scale i, contrast $c_j(x, y)$ and structure $s_j(x, y)$ is calculated and at scale M(M = 5) luminance l_M is computed.

As computation of luminance terms in image is most consuming part in SSIM, Fast SSIM can be utilized for computing multiple scales of image.

The Fast SSIM index between x and y is,

$$\mathrm{F-SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\mu_{GxGy} + C_2)}{(\mu_2^x + \mu_2^y + C_1)(\mu_{Gx}^2 + \mu_{Gy}^2 + C_2)} \tag{8}$$

Where,

$$\mu_{Gx} = \frac{1}{N_xN_y}\sum_{x=1}^{N_x}\sum_{y=1}^{N_y}|\nabla x(x,y)| \tag{9}$$

$$\mu_{GxGy} = \frac{1}{N_x N_y} \sum_{x=1}^{N_x} \sum_{y=1}^{N_y} |\nabla x(x, y)||\nabla y(x, y)| \quad (10)$$

$|\nabla x|$ & $|\nabla y|$ are the gradient magnitude values of image x and y.

In image wavelet coefficients, consistent phase changes occur due to some image distortions. This phase shift of the coefficients does not alter the structural information of the image. To overcome the problem, complex wavelet structural similarity (CW-SSIM) is introduced[12] which is used without preprocessing (registration process) an image. It is computationally less expensive. The CW-SSIM is defined as

$$CW - SSIM_j = \frac{2\sum_{i=1}^{N}(|c_{x,i}||c_{y,i}|) + K}{\sum_{i=1}^{N}(|c_{x,i}|^2 + \sum_{i=1}^{N} c_{y,i}{}^2 + K} \cdot \frac{2\sum_{i=1}^{N} c_{x,i}c_{y,i}^* + K}{2\sum_{i=1}^{N}|c_{x,i}c_{y,i}^*| + K} \quad (11)$$

The first term is used to determine the magnitudes of the coefficients and the second term is used to find the consistent changes in phase between c_x and c_y.

The overall performance of an image quality assessment (IQA) can be improved by combining information content weighting with SSIM[13]. This algorithm is termed as Information content weighted SSIM (IW-SSIM).

$$IW - SSIM = \prod_{j=1}^{M} (IW - SSIM_j)^{\beta_j} \quad (12)$$

Where,

$$IW - SSIM_j = \frac{\sum_i w_{j,i} c(x_{j,i}, y_{j,i}) s(x_{j,i}, y_{j,i})}{\sum_i w_{j,i}}; \text{for } j = 1, \ldots, M-1 \quad (13)$$

$$\& \ IW - SSIM_j = \frac{1}{N_j} \sum_i l(x_{j,i}, y_{j,i}) c(x_{j,i}, y_{j,i}) s(x_{j,i}, y_{j,i}); \text{for } j = 1, \ldots, M \quad (14)$$

$W_{j,i}$ be the information content weight computed at the i^{th} spatial location in the j^{th} scale.

A new universal image quality index (UQI) is proposed to model any image distortions by considering the three factors. They are loss of correlation, luminance distortion and contrast distortion[14].

An information fidelity criterion (IFC) of the image is estimated using natural scene statistics (NSS) in connection with distortion models. The reference image is evaluated in wavelet domain and distortion image is expressed as an attenuation and additive Gaussian noise model[15]. Visual information fidelity (VIF) is as similar to IFC but involves the amount of information shared between reference and distorted images[16]. The main disadvantage of VIF is its computational complexity.

Another objective quality assessment, noise quality measure (NQM)[17] is proposed to deal a degraded image which is modeled as an original image subjected to linear frequency and additive noise injection. NQM considers the variations in contrast, image dimensions, variation in local luminance mean and variations in contrast sensitivity with distance. The performance of NQM is better than PSNR.

Medical image classification and retrieval takes a major role in the healthcare environment for diagnosis and therapy plan. As no golden framework is suggested for the retrieval task, many suggestions are provided by various individuals to make the retrieval system to be more efficient. CBIR involves in searching the images from the huge repository for user defined image. As the advantages considered in CBIR are more in number rather than text based query, it is incorporated in medical domain also.

CBIR consists of various steps such as image preprocessing, feature extraction, feature selection and similarity measures. In this work, 12077 training and 1000 testing images which are labeled and classified by the radiologists based on modality, directionality, anatomy and pathophysiology are considered[18].

3. Results and Analysis

From the database, 1762 heterogeneous images acquired by different modality in different conditions and 250 hand images (homogeneous images dataset) posed in different angle are taken for the evaluation of retrieval efficiency.

Features of both homogeneous and heterogeneous dataset images are extracted initially. Totally 199 global and local features such as Tamura (3), Gabor Features (48), Region based features (08), Wavelet moments (64), Global textures (13), Histogram of Gradients (HOG) (81), moment invariants (07) are extracted from an image. The features are extracted from training and testing images separately. To find the closeness between the query (or) testing images, similarity measures are used. In this work, Euclidean distance (L_2), Manhattan or City block (L_1), Relative Deviation, Mahalanobis, Cosine, Chebyshev ($L\infty$) and Spearmen distances are used to find the association between the query and training images[19].

The similar images from the heterogeneous database are retrieved for the given query image (Query Image IRMA code specification - X-ray - overview image - lateral, right - cervical spine - musculoskeletal system – unspecified). The query image and retrieved images are shown in Figure 1.

The time taken by the system to retrieve the above images is 1.1540 second. For this given query image and spearmen distance measure, the calculated measures for precision, recall and F-Score are 0.35, 0.130 and 0.190 respectively. Based on the chosen features and distance measure the retrieved images are ranked sequentially from left to right. But focusing on the retrieved images and its relevancy to the given query image, the feature based retrieval system is not compatible for heterogeneous image database. Because for the given query image, very few images are ranked correctly in the position 1 & 2. In the subsequent ranking the irrelevant images like breast and bone are present. Following to these irrelevant images, the images which are closer to the query image are positioned in further. Hence, the most relevant images are not ranked properly. As this statement is derived from subjective analysis about the retrieved images, objective quality measures are also considered to evaluate the efficiency of feature based retrieval system for heterogeneous images database. The various error metrics between the query image and retrieved images are tabulated in Table 1. From the obtained results of error metrics, some of the values of relevant images are overlapping with the values of non-relevant images. The same methodology is also applied to homogeneous database to evaluate the efficiency of retrieval system. The retrieved images are shown in Figure 2. The precision, recall and F score for the above

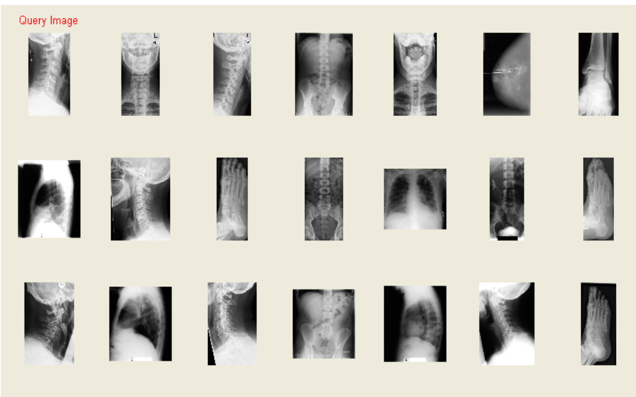

Figure 1. Retrieved images for the query image using Spearmen.

Table 1. Various error metrics and its values of retrieved images from heterogeneous images database.

Image ID	SSIM	MSSIM	CWSSIM	IWSSIM	FSIM	VIF	VIFP	UQI	IFC	NQM	SNR	PSNR	IWPSNR	VSNR	WSNR	MAE	RMSE
741793	0.727	0.333	0.470	0.076	0.699	0.029	0.052	0.046	0.119	−3.036	6.276	10.899	13.082	2.176	6.509	58.201	72.710
741802	0.799	**0.420**	**0.568**	**0.150**	**0.729**	**0.031**	**0.070**	**0.089**	**0.140**	**−1.583**	**7.562**	**12.185**	**13.757**	**4.777**	**7.806**	**45.658**	**62.702**
10401	0.715	0.430	0.478	0.074	0.737	0.023	0.076	0.092	0.112	−1.277	6.748	11.371	15.091	9.056	6.960	55.285	68.864
741794	**0.747**	**0.314**	**0.411**	**0.086**	**0.681**	**0.026**	**0.053**	**0.054**	**0.105**	**−3.336**	**6.223**	**10.845**	**13.147**	**4.605**	**6.453**	**54.506**	**73.161**
9213	0.436	0.309	0.513	0.179	0.757	0.016	0.028	−0.005	0.097	−2.795	2.526	7.148	14.498	3.683	2.716	91.080	111.976
370828	0.626	0.487	0.572	0.183	0.719	0.023	0.071	0.079	0.119	−1.393	6.468	11.091	14.121	4.840	6.671	54.812	71.119
4753	0.490	0.319	0.362	0.145	0.686	0.022	0.048	0.024	0.110	−5.920	3.983	8.606	12.027	1.424	4.177	77.532	94.680
741840	**0.687**	**0.336**	**0.395**	**0.148**	**0.693**	**0.024**	**0.048**	**0.042**	**0.103**	**−4.011**	**4.925**	**9.548**	**12.995**	**1.932**	**5.146**	**64.404**	**84.948**
190791	0.720	0.277	0.410	0.141	0.710	0.027	0.050	0.036	0.122	−2.969	6.526	11.148	14.271	3.517	6.734	54.815	70.651
10571	0.680	0.274	0.385	0.128	0.734	0.029	0.037	0.025	0.118	−1.910	5.439	10.062	13.609	6.056	5.649	65.972	80.065
3447	0.688	0.372	0.411	0.101	0.753	0.021	0.064	0.052	0.105	−2.337	6.975	11.597	14.721	5.253	7.186	52.761	67.092
2032	0.559	0.291	0.447	0.154	0.713	0.028	0.045	0.049	0.122	−3.809	4.266	8.889	12.669	6.213	4.474	72.506	91.646
190795	0.715	0.358	0.495	0.153	0.725	0.025	0.050	0.050	0.120	−0.731	5.919	10.542	15.551	6.366	6.115	59.854	75.761
741787	**0.716**	**0.386**	**0.459**	**0.142**	**0.705**	**0.027**	**0.048**	**0.051**	**0.116**	**−3.080**	**7.256**	**11.879**	**11.968**	**3.355**	**7.497**	**48.993**	**64.951**
5352	0.641	0.295	0.425	0.119	0.721	0.021	0.063	0.061	0.111	−3.314	6.415	11.037	13.276	5.509	6.621	53.350	71.560
741799	**0.651**	**0.230**	**0.603**	**0.068**	**0.678**	**0.024**	**0.037**	**0.016**	**0.107**	**−5.898**	**4.509**	**9.131**	**11.584**	**0.767**	**4.737**	**69.918**	**89.118**
18483	0.726	0.419	0.509	0.193	0.737	0.022	0.076	0.086	0.106	−1.008	6.234	10.857	15.616	8.584	6.453	56.068	73.062
5455	0.507	0.406	0.463	0.223	0.738	0.020	0.055	0.050	0.107	−2.532	4.339	8.962	14.171	6.877	4.546	70.418	90.876
371577	**0.516**	**0.244**	**0.472**	**0.111**	**0.712**	**0.018**	**0.029**	**0.009**	**0.099**	**−6.386**	**3.133**	**7.756**	**12.429**	**0.204**	**3.329**	**85.351**	**104.410**
190799	0.525	0.322	0.423	0.121	0.691	0.024	0.049	0.031	0.118	−1.924	4.493	9.116	13.528	7.423	4.708	69.024	89.278

Figure 2. Retrieved images for the query image using relative deviation from homogeneous database.

retrieved images for the given query image (Query Image IRMA code specification X ray – Plain radiography – digital – coronal – poteroanterior – left hand – upper extremity – musculoskeletal system) are 0.55, 0.13 and 0.21 respectively.

The associated error values with respect to the query image are listed in Table 2. In this table, the images are ordered hierarchically based on its error metrics values irrespective of its original ranking provided by the distance measure of relative deviation. From the obtained results, except the image id 2314, all relevant images are having non overlapping values. But still the ranking offered by the relative deviation based on the feature values are not appropriate. The correlation between each error metrics are listed in the Table 3.

From the correlation values, all HVS based characteristics have strong correlation among them. The error metric values of all retrieved images are plotted in the Figure 3. As the error metric values of relevant and irrelevant images are not deviated much in their range, it is very difficult task for any retrieval system to be more efficient and reliable to user query.

4. Conclusion

As the images retrieved from the heterogeneous images database, the ranking of the relevant images are not ordered significantly. The mathematical and HVS based error values of relevant and irrelevant images are over-lapped each other. If the user likes to retrieve the images from the image database of same category, HVS based evaluation can be used. In order to improve the efficiency of the retrieval system, the salient features of CBIR and Text based retrieval system can be fused.

Table 2. Various error metrics and its values of retrieved images from homogeneous images database.

Image ID	SSIM	MSSIM	CWSSIM	IWSSIM	FSIM	VIF	VIFP	UQI	IFC	NQM	SNR	PSNR	IWPSNR	VSNR	WSNR	MAE	RMSE
2166	**0.79**	**0.38**	**0.73**	**0.20**	**0.76**	**0.07**	**0.16**	**0.18**	**0.41**	**4.57**	**8.97**	**15.64**	**13.94**	**8.99**	**9.18**	**27.66**	**42.11**
2164	**0.82**	**0.51**	**0.82**	**0.33**	**0.74**	**0.09**	**0.18**	**0.22**	**0.46**	**4.15**	**8.17**	**14.84**	**13.37**	**7.45**	**8.40**	**29.78**	**46.19**
2168	**0.78**	**0.55**	**0.73**	**0.38**	**0.72**	**0.09**	**0.16**	**0.23**	**0.42**	**4.31**	**6.77**	**13.44**	**14.30**	**10.72**	**6.96**	**36.27**	**54.27**
2161	**0.80**	**0.56**	**0.78**	**0.38**	**0.72**	**0.10**	**0.17**	**0.24**	**0.49**	**3.54**	**6.48**	**13.15**	**15.93**	**12.06**	**6.68**	**33.30**	**56.09**
4563	**0.75**	**0.32**	**0.72**	**0.21**	**0.74**	**0.05**	**0.07**	**0.04**	**0.29**	**3.32**	**6.45**	**13.13**	**12.80**	**8.46**	**6.67**	**37.56**	**56.26**
2179	**0.77**	**0.48**	**0.69**	**0.34**	**0.72**	**0.08**	**0.13**	**0.20**	**0.38**	**2.94**	**6.43**	**13.10**	**12.87**	**8.61**	**6.63**	**39.46**	**56.44**
2167	**0.83**	**0.64**	**0.65**	**0.49**	**0.75**	**0.12**	**0.21**	**0.31**	**0.54**	**2.68**	**6.16**	**12.83**	**14.82**	**10.46**	**6.38**	**32.73**	**58.21**
2157	**0.71**	**0.42**	**0.64**	**0.19**	**0.70**	**0.07**	**0.10**	**0.11**	**0.32**	**1.94**	**5.80**	**12.47**	**12.59**	**7.55**	**6.00**	**40.15**	**60.68**
2314	0.77	0.37	0.57	0.19	0.75	0.08	0.12	0.14	0.35	2.41	5.57	12.25	14.99	10.56	5.78	40.95	62.27
2172	**0.74**	**0.37**	**0.60**	**0.18**	**0.73**	**0.07**	**0.12**	**0.12**	**0.34**	**1.35**	**5.45**	**12.12**	**15.05**	**7.45**	**5.65**	**41.32**	**63.15**
2160	**0.73**	**0.35**	**0.61**	**0.18**	**0.71**	**0.08**	**0.12**	**0.14**	**0.35**	**2.09**	**4.89**	**11.56**	**13.58**	**9.63**	**5.09**	**44.88**	**67.35**
2360	0.70	0.37	0.53	0.29	0.71	0.07	0.11	0.12	0.36	0.95	4.88	11.56	13.00	6.96	5.08	50.73	67.41
4547	0.72	0.32	0.51	0.18	0.72	0.06	0.05	0.03	0.27	0.89	4.53	11.20	15.34	7.68	4.73	47.34	70.22
4574	0.69	0.28	0.65	0.09	0.71	0.06	0.07	0.04	0.30	0.46	3.85	10.52	11.66	4.88	4.04	49.17	75.92
2148	0.67	0.28	0.57	0.18	0.68	0.06	0.07	0.06	0.27	−0.55	3.34	10.01	12.06	3.42	3.52	54.19	80.51
4586	0.61	0.22	0.40	0.17	0.71	0.05	0.05	0.01	0.25	−1.26	2.62	9.29	14.20	5.14	2.82	65.63	87.46
4590	0.52	0.28	0.60	0.09	0.71	0.05	0.05	0.03	0.26	−1.04	2.19	8.86	12.03	3.77	2.38	63.09	91.90
4440	0.61	0.29	0.57	0.19	0.71	0.06	0.06	0.03	0.27	−0.54	1.31	7.98	14.22	3.50	1.49	76.85	101.72
2171	0.59	0.15	0.56	0.13	0.67	0.05	0.05	0.01	0.24	−2.38	0.95	7.62	10.31	2.24	1.13	74.04	106.02

* The highlighted image ID's are closely relevant to the query image.

Table 3. Correlation between various error metrics

	SSIM	MSSIM	CWSSIM	IWSSIM	FSIM	SNR	PSNR	IWPSNR	VSNR	WSNR	VIF	VIFP	UQI	IFC	NQM
PSNR	0.686	0.800	0.699	0.881	0.97	1	*	0.9944	0.97	1	1	0.99	0.9	0.99	0.959
RMSE	−0.9	−0.865	−0.8153	−0.7868	−0.8	−0.991	−0.74	−0.6652	−0.82	−0.991	−1	−0.8	−1	−0.7	−0.949
MAE	−0.9	−0.855	−0.8201	−0.747	−0.8	−0.967	−0.68	−0.5983	−0.77	−0.967	−1	−0.7	−1	−0.6	−0.927
RMAPE	−0.9	−0.871	−0.7812	−0.7956	−0.8	−0.746	−0.75	−0.6792	−0.84	−0.915	−1	−0.8	−1	−0.7	−0.883
RMSPE	−0.8	−0.703	−0.611	−0.5748	−0.6	−0.887	−0.47	−0.3837	−0.61	−0.887	0	−0.5	−1	−0.4	−0.848

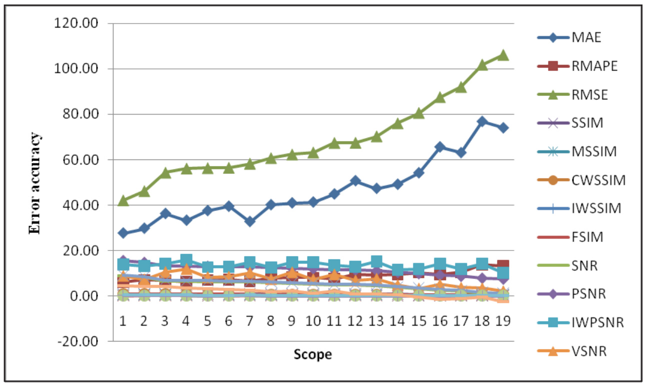

Figure 3. Distribution of error metrics of retrieved images.

5. Acknowledgement

Our sincere gratitude to Prof.Lehman IRMA Group, University of Aachen, Germany for providing the images and IRMA code set to complete this evaluation.

6. References

1. Muller H, Michoux N, Bandon D. A review of content-based image retrieval systems in medicine: clinical benefits and future directions. Int J Med Inform. 2004; 73(1):1–23.
2. Tagare HD, Jaffe C, Duncan J. Medical image databases: a content-based retrieval approach. J Am Med Inform Assoc. 1997; 4(3):184–98.
3. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process, 2004 Apr; 13(4):600–12
4. Pappas TN, Safranek RJ. Perceptual criteria for image quality evaluation. In: Bovik AC, editor. Handbook of image and video processing. Academic Press; 2000 May.
5. Wu HR, Rao KR, editors. Digital video image quality and perceptual coding. Boca Raton, FL: CRC Press; 2005.
6. Girod B. What's wrong with mean-squared error. In: Watson AB, editor. Digital image sand human vision. p. 207–20. Cambridge, MA: MIT Press; 1993.
7. Eskicioglu AM, Fisher PS. Image quality measures and their performance. IEEE Trans Commun, 1995 Dec; 43:2959–65.
8. Wang Z, Bovik AC. Mean squared error: love it or leave it? IEEE Signal Process Mag. 2009 Jan; 98–117.
9. Chandler DM, Hemami SS. VSNR: A wavelet-based visual signal-to-noise ratio for natural images. IEEE Trans Image Process. 2007 Sep; 16(9):2284–98.
10. Wang Z, Bovik AC, Sheikh HR, Simoncelli EP. Image quality assessment: from error visibility to structural similarity. IEEE Trans Image Process. 2004; 13(4):600–12.
11. Wang Z, Simoncelli EP, Bovik AC. Multi-scale structural similarity for image quality assessment. IEEE Asilomar Conf Signals Syst Comput. 2003; 2:1398–402.
12. Sampat MP, Wang Z, Gupta S, Bovik AC, Markey MK. Complex wavelet structural similarity: a new image similarity index. IEEE Trans Image Process. 2009; 18(11):2385–401.
13. Wang Z, Li Q. Information content weighting for perceptual image quality assessment. IEEE Trans Image Process. 2011 May; 20(5):1185–98.
14. Wang Z, Bovik AC. A universal image quality index. IEEE Signal Process Lett. 2002 Mar.
15. Sheikh HR, Bovik AC, De Veciana G. An information fidelity criterion for image quality assessment using natural scene statistics. IEEE Trans Image Process. 2005; 14(12):2117–28.
16. Sheikh HR, Bovik AC. Image information and visual quality. IEEE Trans Image Process. 2006 Feb; 15(2):430–444.
17. Damera-Venkata N, Kite TD, Geisler WS, Evans BL, Bovik AC. Image quality assessment based on a degradation model. IEEE Trans Image Process. 2000; 9(4):636–50.
18. Available from: http://ganymed.imib.rwth-aachen.de/irma/onlinedemos_en.php
19. Vijay J, Bommanna Raja K. Evaluation of Similarity measures in a medical image retrieval system. Int J Appl Eng Res. 2014; 9(21):11039–1105.