Texture Spectral Similarity Criteria Comparison

Michal Havlíček
Institute of Information Theory and Automation, Czech Academy of Sciences
Pod Vodarenskou vezi 4
182 08 Prague, Czechia
havlimi2@utia.cas.cz

Michal Haindl
Institute of Information Theory and Automation, Czech Academy of Sciences
Pod Vodarenskou vezi 4
182 08 Prague, Czechia
haindl@utia.cas.cz
&
Faculty of Management, University of Economics, Jarošovská 1117, Jindřichuv Hradec, Czechia

ABSTRACT

Criteria capable of texture spectral similarity evaluation are presented and compared. From the fifteen evaluated criteria, only four criteria guarantee zero or minimal spectral ranking errors. Such criteria can support texture modeling algorithms by comparing the modeled texture with corresponding synthetic simulations. Another possible application is the development of texture retrieval, classification, or texture acquisition system. These criteria thoroughly test monotonicity and mutual correlation on specifically designed extensive monotonously degrading experiments.

Keywords

Texture Comparison, Texture Modeling, Texture Retrieval, Texture Classification, Texture Acquisition

1 INTRODUCTION

An automatic texture comparison represents a significant but not completely solved complex problem [Hai14]. Such a method would be advantageous to support texture model development where a comparison of the original acquired and to be modeled texture with synthesized or reconstructed ones would help with the optimal model parameter set. There are other possible applications, such as texture database retrieval or texture classification or segmentation, etc. Although there already exist approaches for these tasks, e.g., [Har73, Gal75, Law80, Wys82, Man96, Oja02, Hai06], etc., they do not rank textures according to their visual similarity. Moreover, most methods are limited to mono-spectral textures, a notable disadvantage as color is the most significant visual feature [Hav19].

The psycho-physical evaluations [Hai12], i.e., quality assessments performed by humans, currently represent the only reliable alternative. Methods of this type require both time-demanding experiment design setup and performing, rigorously defined and controlled conditions, and a representative collection of testers, i.e., a sufficient number of individuals, ideally from the general public, naive concerning the goal and design of the experiment. Therefore such experiments are highly impractical and generally demanding, and they cannot be performed on a daily base, on demand, or even in real-time. These experiments are also impracticable in the case of hyper-spectral textures, as not all spectra can be visualized simultaneously due to the limited trichromatic nature of the human perception system.

The criteria mentioned and compared in this paper are intended for the spectral texture composition comparison, i.e., for a specific subset of the general texture comparison problem. The textures are compared as independent sets of pixels where the pixel are treated as vectors of real vector space while the positions of the pixels in the textures are not considered. Texture spectral composition comparison deals with the appearance and amount of pixels that occur in only one of the compared textures and also with the ratio of occurrences of pixels appearing in both textures.

The objectives of the study are as follows:

- To study the effectiveness of spectral similarity criteria for textural applications.
- To analyze their mutual substitutability.

The rest of the paper is organized as follows. Section 2 briefly reviews existing methods relevant to the texture spectral composition comparison. Section 3 outlines experiments used to compare individual methods presented in section 2. Section 3 presents and comments achieved results. Section 4 summarizes the paper with a discussion.

2 TEXTURE SPECTRAL SIMILARITY CRITERIA

In this section, we briefly survey existing texture spectral composition comparison methods. The straightforward way is to use an n-dimensional (n-D) histogram or local one [Yua15], approximating the spectral texture distribution.

Let A and B are the textures to be compared. We denote by a_ρ and b_ρ the ρ-th bin of the n-D histogram of the textures A and B, respectively. The range of the histogram multi-index $\rho = \rho_1, \rho_2, \ldots, \rho_n$ depends on a space C, in which the texture is represented, e.g., in the case of the standard 24-bit red, green, and blue (RGB) color space, the range of all three components of the multi-index is an integer from 0 to 255.

The most intuitive way is to compute the n-D histogram block distance, also known as the Manhattan distance or the Minkowski distance:

$$\Delta_q H(A,B) = \left(\sum_{\rho \in C} |a_\rho - b_\rho|^q \right)^{1/q} , \qquad (1)$$

with $q = 1$ (histogram difference), $q = 2$ (Euclidean distance of histograms), $0 < q < 1$ (fractional dissimilarity of histograms) representing the most used variants. A special case is the maximum distance also called Chebyshev distance or chessboard distance:

$$\Delta_\infty H(A,B) = \sum_{\rho \in C} \max\{|a_{\rho_1} - b_{\rho_1}|, \ldots, |a_{\rho_n} - b_{\rho_n}|\} \ . \qquad (2)$$

Several other possibilities exist for n-D histogram comparison, such as the histogram intersection [Swa91]:

$$\cap H(A,B) = 1 - \frac{\sum_{\rho \in C} \min\left\{a_\rho, b_\rho\right\}}{\sum_{\rho \in C} b_\rho} \ , \qquad (3)$$

the squared chord [Kok03]:

$$d_{sc}(A,B) = \sum_{\rho \in C} \left(\sqrt{a_\rho} - \sqrt{b_\rho} \right)^2 \ , \qquad (4)$$

and the Canberra metric [Kok03]:

$$d_{can}(A,B) = \sum_{C_0} \frac{|a_\rho - b_\rho|}{a_\rho + b_\rho} \ , \qquad (5)$$

where $C_0 = \left\{\rho : a_\rho + b_\rho \neq 0\right\} \subset C$.

The information-theoretic measures like the Kullback-Leibler divergence [Kul51] can also be used:

$$KL(A,B) = \sum_{C^0} a_\rho \log \frac{a_\rho}{b_\rho} \ , \qquad (6)$$

with $C^0 = \{\rho : a_\rho b_\rho \neq 0\} \subset C$, or the Jeffrey divergence:

$$J(A,B) = \sum_{C^J} a_\rho \log \frac{2a_\rho}{a_\rho + b_\rho} + b_\rho \log \frac{2b_\rho}{a_\rho + b_\rho} \ , \qquad (7)$$

can be also considerecan also be considered for n-D histogram comparison as well as a measure based on χ^2 statistic [Zha03]:

$$\chi^2(A,B) = \sum_{C_0} \frac{2\left(a_\rho - \frac{a_\rho + b_\rho}{2}\right)^2}{a_\rho + b_\rho} \ . \qquad (8)$$

The generalized color moments (GCM) [Min98] can also be useful for texture spectral composition comparison problems. The original definition of the GCM of the $(p+q)$-th order and the $(\alpha+\beta+\gamma)$-th degree is:

$$\Delta GCM_{pq}^{\alpha\beta\gamma}(A,B) = \int\int_{\langle A \rangle} r_1^p r_2^q \,[Y^A_{r_1,r_2,1}]^\alpha\, [Y^A_{r_1,r_2,2}]^\beta\, [Y^A_{r_1,r_2,3}]^\gamma dr_1 dr_2 - \int\int_{\langle B \rangle} r_1^p r_2^q \,[Y^B_{r_1,r_2,1}]^\alpha\, [Y^B_{r_1,r_2,2}]^\beta\, [Y^B_{r_1,r_2,3}]^\gamma dr_1 dr_2 \ , \qquad (9)$$

where $[r_1, r_2] \in \langle A \rangle$ represents planar coordinates of the texture pixel Y^A_r, $Y^A_{r_1,r_2,i}$ denotes a pixel intensity in the i-th spectral channel of the texture A, similarly $Y^B_{r_1,r_2,i}$ where $[r_1, r_2] \in \langle B \rangle$. GCM can be easily redefined for an arbitrary number of spectral channels. The terms r_1^p and r_2^q are meaningless in the case of texture spectral composition comparison, and therefore both are put equal to one, using GCMs for which $p = q = 0$ holds. Moreover, it has been observed that the best results are achieved if $\alpha = \beta = \gamma$, specifically using GCMs for $\alpha = \beta = \gamma < 4$ [Hav19].

Another possibility for texture spectral composition comparison represents cosine-function-based dissimilarity, which computes an angle between two vectors.

Figure 1: Textures used for experiments.

Both A, B must have the same number of pixels, a significant drawback of this criterion. This criterion is the only one mentioned in this article suffering from this. All intensity values of corresponding texture spectral channels of all pixels of the textures are arranged into vectors $\vec{A}$ and $\vec{B}$ and the difference is computed as [Zha03]:

$$d_{cos}(A,B) = \frac{\vec{A}^T \vec{B}}{|\vec{A}|\,|\vec{B}|} \quad . \tag{10}$$

Various set-theoretic measures can be considered as criteria as well. Let sets $\mathscr{A}$ and $\mathscr{B}$ denotes the set of unique multi-dimensional vectors representing pixels occurring in the texture A and B, respectively. Criteria can be based on methods developed for comparing the similarity and diversity of the sample sets, such as the Jaccard index [Jac01]:

$$JI(A,B) = \frac{|\mathscr{A} \cap \mathscr{B}|}{|\mathscr{A} \cup \mathscr{B}|} \quad , \tag{11}$$

or the Sørensen-Dice index [Dic45]:

$$SDI(A,B) = \frac{2\,|\mathscr{A} \cap \mathscr{B}|}{|\mathscr{A}| + |\mathscr{B}|} \quad . \tag{12}$$

JI and SDI are equivalent in the sense that given a value for SDI, one can calculate the respective JI value and vice versa.

Alternative to the existing methods may be a modified criterion developed for the texture comparison as the spectral texture composition comparison is its exceptional case. It is possible to remove structure term from the structural similarity metric (SSIM) [Wan04] and define reduced SSIM [Hav16]:

$$rSSIM(A,B) = \frac{1}{\sharp\{r_3\}} \sum_{\forall r_3} \frac{2\mu_{A,r_3}\mu_{B,r_3}}{\mu_{A,r_3}^2 + \mu_{B,r_3}^2} \, \frac{2\sigma_{A,r_3}\sigma_{B,r_3}}{\sigma_{A,r_3}^2 + \sigma_{B,r_3}^2} \quad , \tag{13}$$

where $\sharp\{r_3\}$ is the spectral index cardinality, i.e., the number of spectral channels, μ_{A,r_3} is the mean of r_3-th spectral plane of A and σ_{A,r_3} is the standard deviation of r_3-th spectral plane of A and similarly for μ_{B,r_3} and σ_{B,r_3}.

A very accurate method, the mean exhaustive minimum distance (MEMD), was introduced in [Hav19]. MEMD can be described as the following algorithm. For each pixel from A, the most similar pixel from B is found. This pixel from B can be identified as the most similar to an arbitrary one from A only once. The evaluation ends when all pixels from A have their counterparts in B or all pixels from B are identified as the most similar pixel for an arbitrary one from A. The similarity can

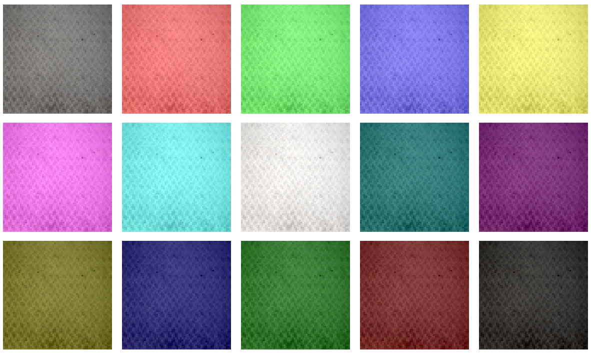

Figure 2: Example of tested texture (top-left) and its the most modified (final) versions obtained during fourteen individual experiments with it.

be expressed by arbitrary metric ρ. The best results were obtained using the maximum metric. The values of metrics are summed and then divided by the number of compared pixels which equals the minimum of the number of pixels in A and the number of pixels in B denoted as M, i.e.:

$$MEMD(A,B) = \frac{1}{M} \sum_{(r_1,r_2)\in\langle A\rangle} \min_{(s_1,s_2)\in\langle U\rangle} \left\{\rho\left(Y^A_{r_1,r_2,\bullet}, Y^B_{s_1,s_2,\bullet}\right)\right\} \quad , \tag{14}$$

where $Y^A_{r_1,r_2,\bullet}$ denotes pixel at $(r_1,r_2)\in\langle A\rangle$, similarly for $Y^B_{s_1,s_2,\bullet}$ but $(s_1,s_2)\in\langle U\rangle$, where $\langle U\rangle$ represents the set of the planar coordinates of the pixels from B not identified as the most similar pixel for the pixels from A evaluated before the pixel at (r_1,r_2).

This criterion was optimized by applying a quicksort sorting algorithm on input data [Hav21]. It is also possible to decrease evaluating time by not including pixels with the exact location in both compared textures if the intensity values are the same in both textures in the corresponding spectral channels. This optimization is meaningful when both textures have the same size, and their difference is expected in the number of pixels, which is significantly smaller than the texture size. It is optional for the locations of such pixels to be known in advance.

3 COMPARISON

All criteria mentioned in the previous section have been extensively tested on precisely defined experiments. The basic idea was to gradually modify the original texture (Figure 1) to resemble the original texture steadily less. The criterion should be able to track these changes to rate the more modified versions of the original texture as less similar to the original. The evaluation error is the ratio of the number of such violations of the assumed monotony to the number of modified versions of the original texture. It should be possible to create modifications that are detectable by the criterion but imperceptible to the human observer. Criteria that can detect even such changes are another advantage over psychophysical experiments and also have possible practical use in areas where maximum accuracy higher than that achievable by a human observer is welcome [Lac22]. Based on these requirements, adjustments were proposed to add or subtract the minimum possible value to all intensity values in selected spectral channels for all texture pixels at once, e.g., Figures 2,3. So that in the case of RGB color space, it is possible to modify data in a single channel, in two channels at the same time, or in all channels at the same time resulting in 14 experiments. In the case of used RGB color space, the minimum possible value that can be added or subtracted equals one, and the adding or subtracting is stopped when maximum, i.e., 255, or minimum, i.e., 0, respectively, is reached

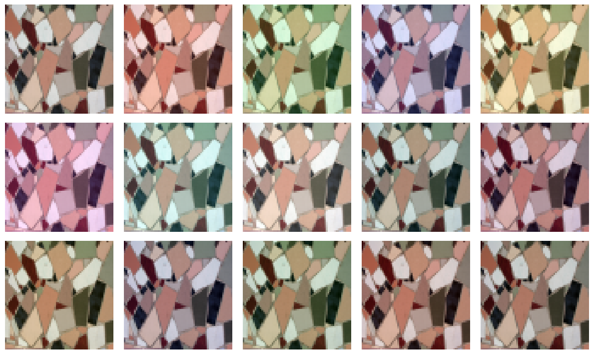

Figure 3: Example of tested texture (top-left) and its the most modified (final) versions obtained during fourteen individual experiments with it.

Criterion	Error [%]	Rank
$\Delta_1 H(\cdot)$	57.0	8
$\Delta_2 H(\cdot)$	57.1	9
$\Delta_{0.5} H(\cdot)$	56.7	7
$\cap H(\cdot)$	57.0	8
$d_{sc}(\cdot)$	56.3	4
$d_{can}(\cdot)$	56.6	6
$KL(\cdot)$	73.1	11
$J(\cdot)$	69.5	10
$\chi^2(\cdot)$	56.4	5
$\Delta GCM^{111}_{00}(\cdot)$	0.0	1
$d_{cos}(\cdot)$	1.1	2
$JI(\cdot)$	48.8	3
$SDI(\cdot)$	59.6	10
$rSSIM(\cdot)$	0.0	1
$MEMD(\cdot)$	0.0	1

Table 1: Average error over all experiments and all textures for individual criteria and corresponding ranks.

criterion	$\Delta_1 H$	$\Delta_2 H$	$\Delta_{0.5} H$	$\cap H$	d_{sc}	d_{can}	KL	J	χ^2	ΔGCM_{00}^{111}	d_{cos}	JI	SDI	$rSSIM$
$\Delta_1 H$														
$\Delta_2 H$	0,99													
$\Delta_{0.5} H$	0,99	0,96												
$\cap H$	1,00	0,99	0,99											
d_{sc}	0,99	0,97	1,00	0,99										
d_{can}	-0,99	0,96	1,00	0,99	1,00									
KL	-0,56	-0,44	-0,68	-0,56	-0,64	-0,66								
J	0,71	-0,61	-0,81	-0,71	-0,78	-0,79	0,97							
χ^2	1,00	0,98	0,99	1,00	1,00	1,00	-0,60	-0,75						
ΔGCM_{00}^{111}	-0,66	-0,60	-0,72	-0,66	-0,69	-0,72	0,71	0,73	-0,68					
d_{cos}	-0,72	-0,65	-0,78	-0,72	-0,75	-0,78	0,76	0,78	-0,73	0,99				
JI	-0,99	-1,00	-0,96	-0,99	-0,97	-0,96	0,44	0,61	-0,98	0,61	0,66			
SDI	-0,87	-0,82	-0,92	-0,87	-0,90	-0,92	0,78	0,85	-0,88	0,90	0,94	0,82		
$rSSIM$	-0,68	-0,61	-0,74	-0,68	-0,71	-0,74	0,74	0,76	-0,69	1,00	1,00	0,62	0,92	
$MEMD$	0,76	0,70	0,81	0,76	0,79	0,81	-0,75	-0,79	0,77	-0,99	-1,00	-0,71	-0,95	-0,99

Table 2: The color criteria Pearson correlation over all 161 materials.

for any intensity value of any pixel. This additional requirement is introduced as a prevention against data overflow or underflow, which could lead to a distortion of the results in the sense that increasing dissimilarity from the original texture would no longer be guaranteed. The number of textures generated by gradually modifying the original texture differs for each texture and depends on the values of the pixel intensities in the original texture. Examples of generated textures to compare with the original can be seen in Figures 2,3.

One hundred sixty-one color textures with resolution 64×64 saved as 24-bit RGB portable network graphics (PNG) files were used as the original textures covering a wide range of natural and artificial materials. Textures were obtained from accessible texture databases [1] [2], and they are shown in Figure 1.

4 RESULTS

Input data used in the experiments described in the previous section led to 78 647 texture-to-texture comparisons for each tested criterion. Achieved results are presented in Table 1. There is an average error over all experiments, and all textures and corresponding ranks are presented for all tested criteria. It is clear from these results that although the tested criteria seem to be theoretically used for texture spectral composition comparison, they rather fail in this task. All histogram-based criteria and set-theoretic measure-based ones reach an error rate of around 50.0% and an even significantly higher error rate in the case of information-theoretic measure-based criteria. One of the reasons might be that our degradation experiments modify non-linearly histograms, often in unpredictable manners, while individual pixels are distorted linearly. But many realist image degradations can be approximated using linear pixel modifications. On the other hand, four criteria meet the requirements for a credible method for texture spectral composition comparison as their error rate is 1.1% (d_{cos}) or even 0.0% (ΔGCM_{00}^{111}, $rSSIM$ and $MEMD$). The target of our paper is not to compare textures. Thus we do not consider here any geometric transformations.

Table 2 illustrates Pearson correlation between all pairs of criteria. The best criteria $(MEMD, rSSIM, \Delta GCM_{00}^{111}, d_{cos})$ are mutually highly correlated ($rSSIM \times \Delta GCM_{00}^{111}$, $rSSIM \times d_{cos}$, $MEMD \times \Delta GCM_{00}^{111}$, $MEMD \times d_{cos}$, $MEMD \times rSSIM$). Similarly, the histogram criteria $(\Delta_1 H(\cdot), \Delta_2 H(\cdot), \Delta_{0.5} H(\cdot), \cap H(\cdot))$, the squared chord $d_{sc}(\cdot)$, and χ^2 are also correlated.

The highly correlated criteria are thus mutually interchangeable.

5 CONCLUSIONS

We properly tested criteria potentially useful for texture spectral composition comparison and demonstrated their suitability in a specially designed experiment. The texture spectral composition comparison represents a partial solution for assessing the textures' quality. Although the criteria do not consider the location of the pixels in the textures, they can help in numerous texture analysis or synthesis applications. The best three criteria - MEMD, generalized color moments, and our reduced structural similarity metric perform with zero spectral ranking errors, while the cosine criterion has a tiny error only. These criteria can be used mainly as a reliable, fully automatic alternative to psychophysical experiments, which are more

[1] texturer.com
[2] mayang.com

impractical due to their cost and strict demands on design setup, conditions control, human resources, and time. Additionally, psychophysical experiments are restricted to visualization of the maximum of 3-D data due to the limited trichromatic nature of human vision, while the MEMD criterion has no upper limit for possible spectral bands.

ACKNOWLEDGMENTS

The Czech Science Foundation project GAČR 19-12340S supported this research.

6 REFERENCES

[Dic45] Dice, L.R. Measures of the amount of ecologic association between species. Ecology, 26, (3), pp.297-302, 1945.

[Gal75] Galloway, M.M. Texture analysis using gray level run lengths. Comput. Graph. Image Process., 4, (2), pp.172-179, 1975.

[Hai12] Haindl, M., Filip, J. Visual texture. in Singh, S., Kang, S.B. (Eds.): Advances in computer vision and pattern recognition, Springer-Verlag London, London, 2012.

[Hai14] Haindl, M., Kudělka, M. Texture fidelity benchmark. in Computational Intelligence for Multimedia Understanding (IWCIM), 2014 International Workshop on, IEEE Computer Society CPS, Los Alamitos, pp.1-5, 2014.

[Hai06] Haindl, M., Mikeš, S. Unsupervised texture segmentation using multispectral modelling approach. Proc. 18th Int. Conf. Pattern Recognition ICPR 2006, vol. II, pp.203-206, 2006.

[Har73] Haralick, R.M., Shanmugam, K., Dinstein, I. Textural features for image classification. IEEE Trans. Syst. Man Cybern., SMC-3, (6), pp.610-621, 1973.

[Hav16] Havlíček, M., Haindl, M. Texture spectral similarity criteria. Proceedings of the 4th CIE Expert Symposium on Colour and Visual Appearance, Commission Internationale de L'Eclairage CIE Central Bureau, Vienna, pp.147-154, 2016.

[Hav19] Havlíček, M., Haindl, M. Texture spectral similarity criteria. IET Image Processing, 13(11), pp.1998-2007, 2019.

[Hav21] Havlíček, M., Haindl, M. Optimized texture spectral similarity criteria. In Advances in Computational Collective Intelligence: 13th International Conference, ICCCI 2021, Kallithea, Rhodes, Greece, September 29-October 1, 2021, Proceedings 13, Springer International Publishing, pp.644-655, 2021.

[Jac01] Jaccard, P. Etude comparative de la distribution florale dans une portion des Alpes et du Jura. Impr. Corbaz, Basel, Switzerland, 1901.

[Kok03] Kokare, M., Chatterji, B., Biswas, P. Comparison of similarity metrics for texture image retrieval. In: TENCON 2003. Conference on Convergent Technologies for the Asia-Pacific Region. Vol. 2. IEEE, pp.571-575, 2003.

[Kul51] Kullback, S., Leibler, R.A. On information and sufficiency. Annals of Mathematical Statistics. 22 (1), pp. 79-86, doi:10.1214/aoms/1177729694, JSTOR 2236703. MR 0039968, 1951.

[Lac22] Lachaud, G., Conde-Cespedes, P., Trocan, M. Patch Selection for Melanoma Classification. In Computational Collective Intelligence: 14th International Conference, ICCCI 2022, Hammamet, Tunisia, September 28-30, 2022, Proceedings, Cham: Springer International Publishing, 2022.

[Law80] Laws, K.I. Rapid texture identification. Proc. SPIE Conf. Image Processing for Missile Guidance, San Diego, USA, pp.376-380, 1980.

[Man96] Manjunath, B.S., Ma, W.Y. Texture features for browsing and retrieval of image data. IEEE Trans. Pattern Anal. Mach. Intell., 18, (8), pp.837-842, 1996.

[Min98] Mindru, F., Moons, T., Gool, L. V. Color-based moment invariants for viewpoint and illumination independent recognition of planar color patterns. In: Illumination Independent Recognition of Planar Color Patterns, Proceedings ICAPR'98. pp.113-122, 1998.

[Oja02] Ojala, T., Pietikäinen, M., Mäenpää, T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. IEEE Trans. Pattern Anal. Mach. Intell., 24, (7), pp.971-987, 2002.

[Swa91] Swain, M. J., Ballard, D. H. Color indexing. International Journal of Computer Vision 7 (1), pp.11-32, 1991.

[Wan04] Wang, Z., Bovik, A.C., Sheikh, H.R., et al. Image quality assessment: from error visibility to structural similarity. IEEE Trans. Image Process., 13, (4), pp.600-612, 2004.

[Wys82] Wyszecki, G., Stiles, W.S. Color science, vol. 8, Wiley, New York, 1982.

[Yua15] Yuan, J., Wang, D., Cheriyadat, A.M. Factorization-based texture segmentation. IEEE Trans. Image Process., 24, (11), pp.3488-3497, 2015.

[Zha03] Zhang, D., Lu, G. Evaluation of similarity measurement for image retrieval. In: Neural Networks and Signal Processing, Proceedings of the 2003 International Conference on. Vol. 2. IEEE, pp.928-931, 2003.