Robust Image Retrieval Based on Mixture Modeling of Weighted Spatio-color Information

Maria Luszczkiewicz-Piatek[1] and Bogdan Smolka[2]

[1] University of Lodz, Faculty of Mathematics and Computer Science, Department of Applied Computer Science, Banacha 22, 90-238 Lodz, Poland
mluszczkiewicz@math.uni.lodz.pl
[2] Silesian University of Technology, Department of Automatic Control, Akademicka 16 Str., 44-100 Gliwice, Poland
bogdan.smolka@polsl.pl

Abstract. In this paper we propose a novel approach to color image retrieval. Color information is modeled using Gaussian mixtures and incorporates the information on the spatial distribution of the color image pixels utilizing the Dijkstra algorithm. The proposed algorithm has high indexing performance and operates on model of low dimensionality. Thus, the proposed method needs only the adjustment of Gaussian Mixture Model parameters for efficient color image retrieval. The proposed method is extensively tested on Corel and Wang dataset. The results demonstrate that proposed framework is more efficient than other methods when images are subjected to lossy coding such as JPEG method.

1 Introduction

In the recent years, the amount of digital images available has grown rapidly due to affordable digital cameras and high-speed Internet connections. Those factors have created a convenient way to generate and publish everyday enormous amount of visual content available worldwide by a growing number of users. Thus, there is a huge demand for image management tools designed for purposes of storing, browsing and searching in large multimedia databases.

In Content-Based Image Retrieval (CBIR) systems, the proper image description is a very important element assessing the similarities among images. Thus, image descriptors can be classified depending on the image property such as color or texture. In CBIR systems, the searching process is based on user given queries, which is often an image, represented by its properties encoded into features vectors, based on concepts of the i.e. color histogram, the color coherence vectors, the color co-occurrence matrix, vector quantization, and then compared in order to find similar images. Also the spatial organization of colors has been explored in form of spatial statistics between color pixels, such correlograms or some filter responses [1,2,3].

Let us note that due to problems related to the high dimensionality of data [5], it is preferable to employ fewer features to accurately represent image information and reduce computational cost, without deteriorating discriminative capability. This approach attracted great attention in recent years, [6].

R.S. Choraś (ed.), *Image Processing & Communications Challenges 6*,
Advances in Intelligent Systems and Computing 313, DOI: 10.1007/978-3-319-10662-5_11

This paper addresses the problem of similar image retrieval, when query image is given, on the basis of the image signature composed of Gaussian Mixture Model parameters. Although several proposed image retrieval techniques utilize the *Gaussian Mixture Model* (*GMM*) as color distribution descriptor [7,8,9], the aspect of the distortions caused by the lossy compression was not taken into account. The spatial arrangement of the color regions present in the analyzed image is incorporated during the color information modelling in form of color homogeneity information evaluated using Dijkstra algorithm, [10]. Due to the smoothing properties of the Gaussian mixtures, this approach is robust to distortions introduced by lossy codding, i.e. *JPEG* compression scheme. The paper is organized as follows. The details of the proposed technique are described in Section 2, it also focuses on the evaluation of spatial arrangement of the color pixels within the scene depicted in the image using the Dijkstra algorithm. Next Section 3 presents the experimental setup of the retrieval scheme based on the proposed solution. The comparison of the experimental results for various compression ratios is also presented in Section 4. Finally, Section 5 concludes the concepts and results presented in this paper.

2 Gaussian Mixture Modeling

The first and very important decision concerning the color image data modeling using any technique is the choice of the color space suitable for the retrieval experiments [11]. In this paper we are using the $CIE\ La^*b^*$ color space. The first step in applying the proposed methodology is to construct the histogram $H(x, y)$ in the $a - b$ chromaticity space defined as $H(x, y) = N^{-1}\#\{a_{i,j} = x; b_{i,j} = y$, where $H(x, y)$ denotes a specified bin of a two-dimensional histogram with a the component equal to x and b component equal to y, the symbol $\#$ denotes the number of samples in a bin and N is the number of color image pixels. The next stage of the presented technique is the modeling of the color histogram using the $Gaussian\ Mixture\ Model$ (GMM) and utilizing the Expectation-Maximization (EM) algorithm for the model parameters estimation as described in details in [12], using following formulae:

$$\alpha_m^{k+1} = N^{-1} \sum_{i=1}^{N} p(m|x_i, \Theta^k), \quad \mu_m^{k+1} = \frac{\sum_{i=1}^{N} x_i \cdot p(m|x_i, \Theta^k)}{\sum_{i=1}^{N} p(m|x_i, \Theta^k)}, \tag{1}$$

$$\upsilon_m^{k+1} = \frac{\sum_{i=1}^{N} p(m|x_i, \Theta^k)(x_i - \mu_m^{k+1})(x_i - \mu_m^{k+1})^T}{\sum_{i=1}^{N} p(m|x_i, \Theta^k)} \quad , \tag{2}$$

where μ and υ denote the mean and variance, m is the index of the model component and k is the iteration number. The E (Expectation) and M (Maximization) steps are performed simultaneously, according to (1) and (2) and in each iteration, and the input data we use parameters obtained in the previous one. The main idea of the application of the GMM technique lies in the highly desirable properties of this approach. The inherent feature of the GMM enables to approximate the distorted color histogram of the color image subjected to lossy compression, which is obtained through limited model complexity (7 components) and number of iterations (75) of E-M algorithm, as shown in [15,14,13,16]. The lossy compression causes a loss of color information resulting in discontinuities in the chromaticity histogram. The proposed methodology counteracts this effect by smoothing the histogram in order to reconstruct its original shape, which is a basis for the effective image retrieval, which lack other approaches based on mixture modeling [17].

Thus, in order to reflect the differences in image content and also to improve the retrieval efficiency of the proposed retrieval scheme, the concept of using *Dijkstra algorithm* [10], was incorporated. Its use provides a possibility to emphasize in the $a-b$ color histogram all the pixels belonging to homogenous color areas. In general, its construction reflects topological *distance* and color *similarity* of image pixels, evaluated using $CIEDE2000$ color similarity measure [18], represented by *optimal paths* between central pixel and each of its neighbors in processing window.

Using the concept of *Dijkstra algorithm*, each pixel of the original image is taken into the $a-b$ histogram with the weight evaluated on the basis of the sum of the costs of the "cheapest" paths to pixels belonging to processing window W, where cost of visiting neighboring pixels is calculated as difference of their colors using $CIEDE2000$ (Fig. 1). Thus, chromaticity histogram is in fact a histogram of weights associated to color pixels. Each bin value is the sum of the weights of the pixels of the particular $a-b$ values. The idea of evaluating color region homogeneity is based on treating processing window W (assigned as $h_1 = 10\%$ and $h_2 = 10\%$ of image height and width respectively) centered at pixel $y(i,j)$ and consisted of K pixels, as a directed graph.

The Dijkstra algorithm creates the path of the lowest total cost from the center pixel $y(i,j)$ of the processing window W to the pixel $y(i,j)^{(k)} \epsilon W$ and assigns the optimal cost $p(i,j)^{(k)}$ The weight associated with pixel $y(i,j)$, is based on the average of all optimal path cost evaluated for pixels of processing window W: $w(i,j) = \exp\left(-\frac{\sum_{k=1}^{K} p(y(i,j)^{(k)})^2}{h_1 \cdot h_2}\right)$. Thus, $w(i,j)$ can be treated as similarity measure between central pixel $y(i,j)$ and it surrounding $y(i,j)^{(k)} \epsilon W$. Let us note that, the more homogenous the pixel neighborhood, the lower is the the sum of costs of the optimal paths to all window pixels, and thus the larger is the weight value $w(i,j)$. In case of regions presenting significant color changes the total cost of optimal paths will be larger due to observed color differences among pixels.

3 Experimental Setup

In order to test the proposed methodology we subjected to lossy coding ($JPEG$) the database of Wang [4] consisting of 1000 natural images, of 10 semantic categories, and Corel dataset of 10 000 color images. For the purposes of the histogram comparisons, several distance or similarity measures were used (denoted

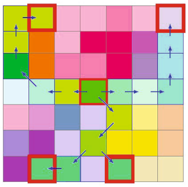

Fig. 1. Exmplary optimum paths with minimum costs for Dijkstra algorithm using $CIEDE2000$ as a color similarity measure

as d) and computed between the histogram of the evaluated image (denoted as H) and the 2D surface generated by the GMM model of its histogram, (denoted as C). Among them are geometric measures treating objects as vectors in a multi-dimensional space and computing the distance between two data points on the basis of pairwise comparisons in each dimension (Minkowski family distances such as L_1 and L_2 and Canberra metric). Information-theoretic measures (such as Jeffrey divergence) derived from Shannon's entropy theory and statistical measures (such as χ^2) compare objets not in pairwise manner but rather as a distributions. Following formulas describe applied similarity measures:

– L_1 (D_{L_1}) : $d(H,C) = \sum_{i=1}^{\eta}\sum_{j=1}^{\eta}|H_{i,j} - C_{i,j}|$,
– L_2 (D_{L_2}): $d(H,C) = \left(\sum_{i=1}^{\eta}\sum_{j=1}^{\eta}|H_{i,j} - C_{i,j}|^2\right)^{\frac{1}{2}}$,
– Canberra: (D_C) $d(H,C) = \left(\sum_{i=1}^{\eta}\sum_{j=1}^{\eta}\frac{|H_{i,j}-C_{i,j}|}{|H_{i,j}|+|C_{i,j}|}\right)$,
– Jeffrey Divergence (D_J): $d(H,C) = \left(\sum_{i=1}^{\eta}\sum_{j=1}^{\eta} H_{i,j}\cdot\log\frac{H_{i,j}}{\eta_{i,j}} + C_{i,j}\cdot\log\frac{C_{i,j}}{\eta_{i,j}}\right)$,
– χ^2 (D_{χ^2}): $d(H,C) = \left(\sum_{i=1}^{\eta}\sum_{j=1}^{\eta}\frac{(H_{i,j}-\eta_{i,j})^2}{\eta_{i,j}}\right)$,

where $\eta_{i,j} = \frac{H_{i,j}+C_{i,j}}{2}$. For the evaluation of the difference between two histograms we also used the the *Earth Mover's Distance* (EMD) similarity measure, [19]. The EMD is based on the assumption that one of the histograms reflects "hills" and the second represents "holes" in the ground of a histogram. The measured distance is defined as a minimum amount of work needed to transform one histogram into the other using a "soil" of the first histogram. As this method operates on signatures and their weights using GMM, we assigned as signature values the *mean* of each component and for the *signature weight* the weighting coefficient of each Gaussian in the model.

4 Evaluation of the Method Efficiency

In order to present retrieval observations, the *Precision* and *Recall* measures are employed. In more details, *Precision* is the fraction of retrieved images that are relevant, while Recall is the fraction of relevant instances that are retrieved. Fig. 2 illustrates the *Precision* – *Recall* plots evaluated for the entire analyzed

Wang database, i.e. each of the database image was used as a query. In these experiments the criterion of the successful retrieval was the membership to the same thematic group as the query, not necessarily sharing the same color palette. Therefore, the main aim of the presented plots and numerical results is to show that proposed technique can produce meaningful results regardlessly to the rate of information loss associated with various compression methods. Fig. 2 compares the retrieval efficiency for several similarity measures. It can be noticed that there is no significant difference between retrieval accuracy for various compression ratios. The evaluated *Precision* values for each of 1000 queries were averaged (for each corresponding *Recall* value) producing the results shown in Fig. 3. The use of *EMD* similarity measure produced the best results, and moreover it offers compact color description contrary to other measures operating on histograms.

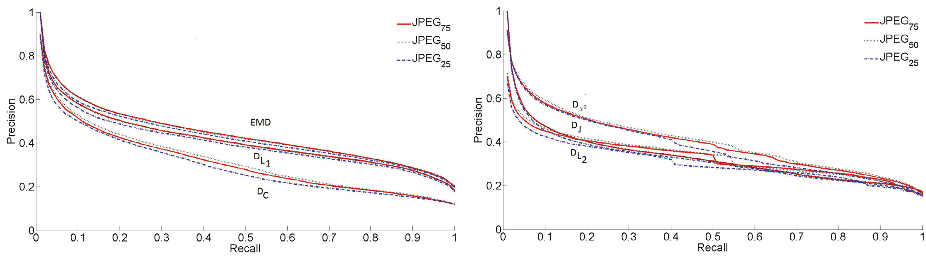

Fig. 2. The $Precision - Recall$ plots evaluated for Wang database of 1000 natural color images subjected to lossy coding using compression rates: 75% ($JPEG_{75}$), 50% ($JPEG_{50}$) and 25% ($JPEG_{25}$) for various similarity measures: D_{L_1}, D_{L_2}, D_C metrics, D_J, D_{χ^2} and EMD

Dataset	EMD	D_C	D_{L_1}	D_{L_2}	D_J	D_{χ^2}
$JPEG_{75}$	0.4125	0.3099	0.4384	0.3302	0.3324	0.3941
$JPEG_{50}$	0.4109	0.3194	0.4370	0.3304	0.3353	0.4030
$JPEG_{25}$	0.4008	0.2948	0.4272	0.3214	0.3062	0.3791

Fig. 3. The average *Precision* values evaluated for Wang database of 1000 natural color images subjected to $JPEG$ lossy coding using compression rates: 75% ($JPEG_{75}$), 50% ($JPEG_{50}$) and 25% ($JPEG_{25}$) for various similarity measures: D_{L_1}, D_{L_2}, D_C metrics, D_J, D_{χ^2} and (EMD)

Let us note, that *Precision* and *Recall* tend to ignore the ranking of retrieved images, e.g. *Precision* of 0.5 indicates that half of the retrieved images is relevant to the given query but without any further information if relevant images are the first half of the retrieved set or are the second half. In order to counteract this, it is advised to measure *Precision* and *Recall* at the specified points e.g. at the answer set of 1 ($\hat{P_1}$), 5 ($\hat{P_5}$), 10 ($\hat{P_{10}}$), and 25 ($\hat{P_{25}}$) images. These measurements (Fig. 4) are summarized to give the better view of the behavior of the analyzed

	EMD				D_C				L_1 (D_{L_1})			
Dataset	$\hat{P}_1$	$\hat{P}_5$	$\hat{P}_{10}$	$\hat{P}_{25}$	$\hat{P}_1$	$\hat{P}_5$	$\hat{P}_{10}$	$\hat{P}_{25}$	$\hat{P}_1$	$\hat{P}_5$	$\hat{P}_{10}$	$\hat{P}_{25}$
$JPEG_{75}$	1	0.6569	0.5810	0.4794	1	0.6569	0.5810	0.4794	1	0.6569	0.5810	0.4794
$JPEG_{50}$	1	0.6531	0.5795	0.4780	1	0.6531	0.5795	0.4780	1	0.6531	0.5795	0.4780
$JPEG_{25}$	1	0.6430	0.5628	0.4647	1	0.6430	0.5628	0.4647	1	0.6430	0.5628	0.4647

	D_{L_2}				D_J				D_{χ^2}			
Dataset	$\hat{P}_1$	$\hat{P}_5$	$\hat{P}_{10}$	$\hat{P}_{25}$	$\hat{P}_1$	$\hat{P}_5$	$\hat{P}_{10}$	$\hat{P}_{25}$	$\hat{P}_1$	$\hat{P}_5$	$\hat{P}_{10}$	$\hat{P}_{25}$
$JPEG_{75}$	1	0.5574	0.4733	0.3814	1	0.5576	0.4716	0.3800	1	0.5482	0.4541	0.3714
$JPEG_{50}$	0.6996	0.5003	0.4496	0.3926	0.7140	0.5161	0.4626	0.4011	0.663	0.4734	0.4246	0.3668
$JPEG_{25}$	0.9097	0.6552	0.5770	0.4781	0.9116	0.6653	0.5919	0.4872	0.8956	0.6480	0.5706	0.4743

Fig. 4. The average *Precision* values evaluated for Wang database of 1000 natural color images subjected to lossy coding using compression rates: 75% ($JPEG_{75}$), 50% ($JPEG_{50}$) and 25% ($JPEG_{25}$) for various similarity measures: D_{L_1}, D_{L_2}, D_C metrics, D_J, D_{χ^2} and EMD at points corresponding to 1, 5, 10 and 25 retrieved images

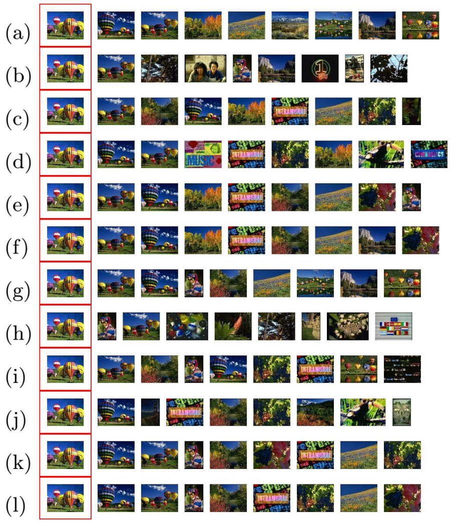

Fig. 5. The comparison of the retrieval results performed for Corel 10k database obtained for various similarity measures: EMD (a,g), *Canberra* (b,h), L_1 (c,i), L_2 (d,j), *Jeffrey divergence* (e,k), χ^2 (f,l). The results were evaluated for images transformed to GIF (a-f) and JPG_{25} (g-l)

retrieval scheme for applied color spaces. The EMD similarity measure produced the best results.

Fig. 4 illustrates the comparison between retrieval results evaluated for exemplary image of the Corel database of 10 000 color images for the various similarity measures for lossy coded versions of original images (25% original file size) and GIF (using dithering) method. It can be noticed that the evaluated results preserve the spatio-chromatic structure of the query image. The L_2 similarity measure yield the worst results, and should not be used for image indexing purposes.

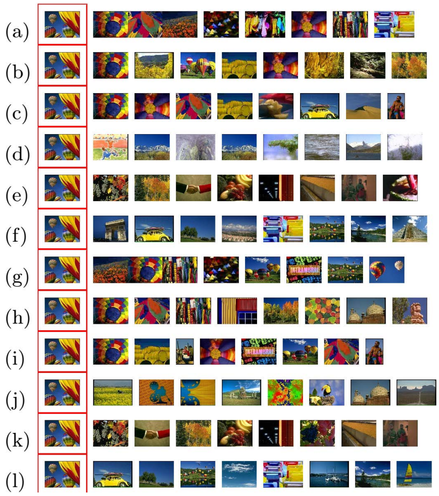

Fig. 6. The comparison of the retrieval results performed for Corel 10k database obtained for various retrieval methods: GMM with Dijkstra weighting applied for the images represented in the $CIE\ La^*b^*$ color space (a,g), CEDD [20] (b,h), FCTH [20] (c,i), MPEG-7 SCD (d,j), MPEG-7 CLD (e,k), Autocorrelogram (f,l). The results were evaluated for images transformed to GIF (a-f) and JPG_{25} (g-l)

The retrieval results evaluated for EMD similarity measure, were compared to those obtained using the well known, image retrieval application img(Rummager) [20]. This system contains several methods, from which the following were chosen: 60 bin histogram (CEDD), 192 bin histogram FCTH, MPEG-7 SCD and MPEG-7 CLD, and Autocorrelogram. Fig. 4 presents the retreval results for exemplary image of the Corel database transformed using GIF and JPG_{25} schemes.

When analyzing the performance of image retrieval techniques in comparison to scheme proposed in this paper, (Fig. 4) it can be noticed that, the color palette query image is not preserved (especially when GIF method was applied), the spatial arrangement the retrieved images is not always taken into account.

5 Conclusion

Many retrieval system rely on the color composition of the analyzed images. Although, this approach seems to be generally correct and effective, one must be aware of the problem of accurately managing the vast amount of visual information. The methods operating on chromaticity histograms could be severely disabled as color palettes of lossy compressed images can differ, providing misleading conclusions. As shown in this paper, such a comparison produces incorrect results when the retrieval process is evaluated not on the original images but on their compressed versions. Thus, there is an urgent need for the evaluation of techniques overcoming that undesirable phenomenon. Such a method is described in this paper.

The main contribution of this work is the adaptation of the Gaussian Mixture Models and the application of the Dijkstra algorithm for the purposes of the distorted chromaticity histogram approximation. The proposed scheme enables the retrieval system not only to take into account the overall image color palette but also to consider the the color dispersion understood as spatial arrangement of image colors. Let us note, that due to the fact that proposed solution preserves regions of homogenous color it can be applied also for retrieval of image of similar background content.

The satisfactory results were achieved independently on the applied compression scheme. Therefore, the presented results proved the hypothesis that the loss of color information caused by lossy coding can be efficiently counteracted, preserving spatial arrangement of colors in an image, providing successful retrieval results.

Acknowledgment. This work has been supported by The National Science Centre under SONATA grant no. UMO-2011/01/D/ST6/04554.

References

1. Ciocca, G., Schettini, L., Cinque, L.: Image Indexing and Retrieval Using Spatial Chromatic Histograms and Signatures. In: Proc. of CGIV, pp. 255–258 (2002)
2. Lambert, P., Harvey, N., Grecu, H.: Image Retrieval Using Spatial Chromatic Histograms. In: Proc. of CGIV, pp. 343–347 (2004)
3. Hartut, T., Gousseau, Y., Schmitt, F.: Adaptive Image Retrieval Based on the Spatial Organization of Colors. Computer Vision and Image Understanding 112, 101–113 (2008)
4. Wang, J.Z., Li, J., Wiederhold, G.: SIMPLIcity: Semantics-Sensitive Integrated Matching for Picture Libraries. IEEE Trans. Patt. Anal. Mach. Intel. 9, 947–963 (2001)
5. Fodor, I.K.: A survey of dimension reduction techniques. Technical Report UCRL- ID-148494, Lawrence Livermore National Laboratory, US Department of Energy (2002)
6. Wu, P., Manjunath, B.S., Shin, H.D.: Dimensionality Reduction for Image Retrieval. In: Proc. of ICIP (2000)
7. Perronnin, F., Liu, Y., Sanchez, J., Poirier, H.: A similarity measure between unordered vector sets with application to image categorization. In: Proc. of CVPR, pp. 3384–3391 (2008)
8. Xing, X., Zhang, Y., Gong, B.: Mixture Model Based Contextual Image Retrieval. In: Proc. of CIVR 2010, pp. 251–258 (2010)
9. Beecks, C., Ivanescu, A.M., Kirchhoff, S., Seidl, T.: Modeling Image Similarity by Gaussian Mixture Models and the Signature Quadratic Form Distance. In: Proc. of ICCV, pp. 1754–1761 (2011)
10. Dijkstra, E.W.: A note on two problems in connexion with graphs. Numerische Mathematik 1, 269–271 (1959)
11. Łuszczkiewicz-Piątek, M.: Which Color Space Should Be Chosen for Robust Color Image Retrieval Based on Mixture Modeling. In: Choras, R.S. (ed.) Image Processing and Communications Challenges 5. AISC, vol. 233, pp. 55–64. Springer, Heidelberg (2014)

12. Bilmes, J.: A Gentle Tutorial on the EM Algorithm and its Application to Parameter Estimation for Gaussian Mixture and Hidden Markov Models. University of Berkeley, ICSI-TR-97-021 (1997)
13. Luszczkiewicz-Piatek, M., Smolka, B.: Selective color image retrieval based on the Gaussian mixture model. In: Blanc-Talon, J., Philips, W., Popescu, D., Scheunders, P., Zemčík, P. (eds.) ACIVS 2012. LNCS, vol. 7517, pp. 431–443. Springer, Heidelberg (2012)
14. Luszczkiewicz, M., Smolka, B.: Application of bilateral filtering and Gaussian mixture modeling for the retrieval of paintings. In: Proc. of ICIP, pp. 77–80 (2009)
15. Luszczkiewicz-Piatek, M., Smolka, B.: Effective color image retrieval based on the Gaussian mixture model. In: Schettini, R., Tominaga, S., Trémeau, A. (eds.) CCIW 2011. LNCS, vol. 6626, pp. 199–213. Springer, Heidelberg (2011)
16. Luszczkiewicz, M., Smolka, B.: Gaussian Mixture Model Based Retrieval Technique for Lossy Compressed Color Images. In: Kamel, M.S., Campilho, A. (eds.) ICIAR 2007. LNCS, vol. 4633, pp. 662–673. Springer, Heidelberg (2007)
17. Avrithis, Y., Kalantidis, Y.: Approximate Gaussian Mixtures for Large Scale Vocabularies. In: Fitzgibbon, A., Lazebnik, S., Perona, P., Sato, Y., Schmid, C. (eds.) ECCV 2012, Part III. LNCS, vol. 7574, pp. 15–28. Springer, Heidelberg (2012)
18. Sharma, G., Wu, W., Dalal, E.N.: The CIEDE2000 color-difference formula: implementation notes, supplementary test data, and mathematical observations. Color Research and Application 30(1), 21–30 (2005)
19. Rubner, Y., Tomasi, C., Guibas, L.J.: The Earth Mover Distance as a Metric for Image Retrieval. International Journal of Computer Vision 40(2), 99–121 (2000)
20. Chatzichristofis, S.A., Boutalis, Y.S., Lux, M.: Img (rummager): An interactive content based image retrieval system. In: Proc. of the 2nd International Workshop on Similarity Search and Applications (SISAP), pp. 151–153. IEEE Computer Society (2009)