A New Similarity Matching Measure
Application to Texture-based Image Retrieval

Noureddine Abbadeni
DMI-Faculty of sciences
University of Sherbrooke
Sherbrooke, Québec, J1K 2R1, Canada
E-mail: abbadeni@dmi.usherb.ca

Abstract

This paper addresses the fundamental issue of similarity in image databases. A new similarity model is introduced based on Gower's coefficient of similarity. This similarity model is flexible and can be declined in several versions: non-weighted, weighted and hierarchical versions. The similarity model is applied on textures by considering two content representation models: the well-known autoregressive model and a perceptual model based on perceptual features such as coarseness and directionality. Experimentations conducted with human subjects, showing interesting results, are presented.

1 Introduction

Among all the fundamental issues in content-based image retrieval, and image databases in general, the problem of similarity matching is particularly important and it constitutes a fundamental operation to carry out image retrieval systems. Similarity can be defined as a mapping function between the set or vector of parameters representing the content of images and a positive real value chosen to quantify the degree of resemblance between the compared images. A similarity model recovers the two following fundamental questions: 1. Choosing the representation model which determines the relevant set of features to represent the content of images; 2. Choosing the theoretical similarity model which indicate in particular how individual or partial similarities on each feature combine to give a global similarity ?

The rest of this papers is organized as follows: In section 2, we give a brief outline on two content representation models for textured images; In section 3, we present the similarity model; In section 4, experimental results are given showing the performance of the similarity model proposed with respect to both of the two representation models; In section 5, a conclusion is given and further investigations related to this work are depicted. Finally, note that related work was omitted in this paper because of lack of space.

2 Similarity model

We present, first, the basic similarity model in which the combination of features to constitute a global similarity is done as an average of each of the individual similarities on each feature. Then we will give two other models to combine individual similarities to form global similarity: the first one is by proposing specific weights for each of the features; the second one is by proposing a hierarchical combination of the features according to the concepts of primary and secondary nature of each feature.

2.1 Basic similarity model

The basic similarity model, the non-weighted model, inspired from Gower's similarity coefficient [7], denoted GS, is defined as follows:

$$GS_{ij} = \frac{\sum_{k=1}^{n} S_{ij}^{(k)}}{\sum_{k=1}^{n} \delta_{ij}^{(k)}} \tag{1}$$

where $S_{ij}^{(k)}$ is the result of the comparison of images i and j on the feature k et $\delta^(k)_{ij}$ represents the possibility to be able to compare images i and j on feature k. $\delta_{ij}^{(k)}$ equals 1 if images i and j can be compared on feature k, 0 if not. In the case that images i and j can be compared on all the considered features, $\sum_{k=1}^{n} \delta_{ij}^{(k)}$ equals N, the number of these features.

GS_{ij}, the global similarity between images i and j, is defined as an average of the similarities on each feature between images i and j.

The quantity $S_{ij}^{(k)}$ can be defined as follows:

$$S_{ij}^{(k)} = 1 - \frac{|x_{ik} - x_{jk}|}{R_k} \tag{2}$$

where R_k represents a normalization factor. $S_{ij}^{(k)}$ equals 1 if images i and j are identical, 0 if they are completely different, a positive value between 0 and 1 if the two images have a certain degree of similarity according to feature k.

Using equation (2) and considering that all features can be compared, the global similarity GS_{ij} between two images i and j can be re-written as follows:

$$GS_{ij} = 1/N \sum_{k=1}^{n} (1 - \frac{|x_{ik} - x_{jk}|}{R_k}) \tag{3}$$

The values taken by different features must be normalized in order to maintain an equal intervalle among all the compared pairs of images since this normalization can be done using the intervalle of values taken by each feature for each of the images compared. This can be done across all the sample or the database of images considered. The intervalle of possible values, denoted R_k for feature k, is computed as $R_k = Max(x_{ik}) - Min(x_{ik})$ where $x_{ik}, i = 1..n$, is the set of values taken by each of the image i of the sample considered for the feature k. Max and Min represent respectively the maximum and the minimum value.

2.2 Weighted similarity model

When features representing the content of an image have not the same importance, which occurs often in practice, it is common to associate different weight to each feature according to its importance. The weighted similarity model can be written as follows:

$$GS_{ij} = \frac{\sum_{k=1}^{n} w_k S_{ij}^{(k)}}{\sum_{k=1}^{n} w_k \delta_{ij}^{(k)}} \tag{4}$$

where w_k represents the weight associated to feature k. Note that w_k is constant, for each feature, for all the images of the sample/database considered.

2.3 Hierarchical similarity model

Another way of resolving the importance to give to a feature compared to another is the use of the concepts of primary features and secondary features. These are important concepts that allow to distinguish between a feature that has a proper existence independently from the existence of other features from a feature that exist dependently from another feature. The concepts of primary and secondary feature can also be defined from the perceptual importance of a feature or another. In this sense, a good similarity measure should never allow secondary features to have the same influence or more than primary features on the global similarity between two images. One manner to take into account this phenomena is to consider the similarity $S_{ij}^{(k)}$ between two images i and j according to primary features k_p. Each similarity on each primary feature is multiplied by the similarity between images i and j according to all of the secondary features k_s, if they exist, associated to primary feature k_p. The hierarchical similarity model can be written as follows:

$$GS_{ij} = \frac{\sum_{k=1}^{m} GS_{ij}^{(k)} S_{ij}^{(k)}}{\sum_{k=1}^{m} \delta_{ij}^{(k)}} \tag{5}$$

where M is the number of primary features at each level of the hierarchy. If a primary feature does not have secondary features associated with it, $GS_{ij}^{(k)} = 1$. The summation is done only on primary features. Similarities on secondary features, when they exist, play a role of variable weights for the primary feature to which they are associated. This hierarchical similarity measure have at least the following desirable properties: 1. It gives more importance to primary features and, most importantly, does not give secondary features much more importance than primary ones; 2. Similarity on secondary features allow to define a function of variable weight for each primary feature that has secondary features associated with it; 3. And finally, note the recursive nature of the hierarchical similarity measure. The so defined hierarchical similarity model gives an original manner to combine individual (or partial) similarities on each feature to form a global similarity using the concepts of primary and secondary features. Finally, note that it is possible to combine both of the weighted and hierarchical models to obtain a hybrid model, weighted and hierarchical at the same time.

3 Representation models

We have considered two content representation models for textured images: a statistical model, the well-known autoregressive model [3] and an empirical perceptual model based on a set of perceptual features namely coarseness, directionality, contrast and busyness [1], [2]. Because of lack of space, we cannot reproduce all the theory behind the two models. We will briefly outline the basic principles of each model. The reader interested in more details on these representation models should consult the mentioned references.

The autoregressive model is well-known and was extensively used to model textured images. Both the separable and non-separable versions of the autoregressive model were experimented with different neighborhoods. The different models were evaluated in two ways: qualitative evaluation by measuring their ability to capture the visual content of textured images and reconstruct images from the estimated parameters of the model and a quantitative evaluation by measuring the mean square error between the original image and the reconstructed one. Experimental results

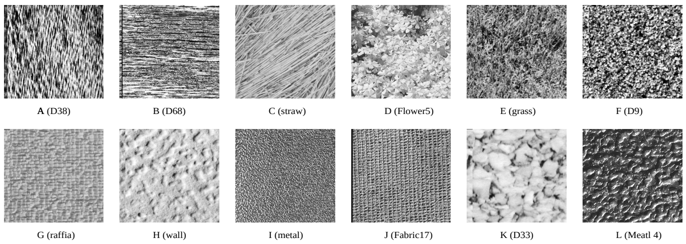

Figure 1: Sample of test images.

showed the superiority of the non-separable (2D) autoregressive model with a neighborhood of size 3 (asymmetric quarter-plan). Thus this is the version of the autoregressive model which is considered in the experimental results [3].

The perceptual model relies on four perceptual features: coarseness, directionality, contrast and busyness. Estimation techniques of these perceptual features was based on the auto-covariance function. Coarseness was estimated as the average number of extrema in the auto-covariance function applied to images. Directionality was estimated as the average number of pixels having the dominant orientation(s). Contrast was estimated as a combination of three parameters: coarseness, the average magnitude of the gradient of the auto-covariance function and the average number of pixels having a significant magnitude (superior on a certain threshold). Busyness was estimated totally based on coarseness. The computational estimations of each of these four perceptual features were evaluated with human judgments based on a psychometric method. This psychometric method is based on the sum of rank values and the Spearman's rank of correlation coefficient. Experimental results showed very interesting results and strong correlation between the computational measures proposed and the human judgments was found. Values of Spearman rank-correlation coefficient, denoted r_s, found are as follows: for coarseness, $r_s = 0.913$; for directionality, $r_s = 0.841$; for contrast, $r_s = 0.755$; and finally, for busyness, $r_s = 0.774$ [1].

4 Experimental results

4.1 Human judgments of similarity

We have conducted psychological experiments with human subjects on similarity of textured images. We have presented twelve textured images and we have asked human subjects to choose, for each of the twelve images, the top three most perceptually similar images. The sample of images used is given in figure (Fig. 1).

Thirty human subjects have participated in the experiments. To obtain one consolidated human judgment of similarity, we have used a psychometric method based on the sum of ranked values which was also used in our earlier work on perceptual features [1] and described in details in [10] and was used also by [4] and [12]. The consolidated human judgment of similarity related to images of figure (Fig. 1) is given in table (Tab. 4.1).

Img	1^{st}	2^{nd}	3^{rd}
A	E	F	B
B	A	E	G
C	B	G	E
D	K	F	H
E	F	A	L
F	E	A, L	I
G	J	I	B
H	K	L	D
I	F	J	E
J	G	I	F, A
K	H	D	L
L	F	H	E

Table 1: Human subjects consolidated judgment of similarity: the top three most similar images to each image in the sample considered.

4.2 The autoregressive model

Textured images content is represented by the autoregressive model. In this case, we can point out that there is an average correspondence between the results obtained by the autoregressive model with respect to human judgments of similarity. However images that have been classified in the three top positions by human subjects obtain, often, a high score of similarity even if they are not classified in the top three positions.

4.3 The perceptual model

4.3.1 Non-weighted combination of perceptual features

Several combinations were experimented. The first combination consider only coarseness and directionality. Coarseness and directionality were found in our earlier work [1] as the most important perceptual features. The second combination consider coarseness, directionality and contrast. The third combination consider coarseness, directionality, contrast and busyness. Results obtained, compared to human rankings of similarity, show that the combination of all the four features gives better results than the other tow combinations.

4.3.2 Weighted combination of perceptual features

The same combinations as in the precedent section are reproduced in this section except that we add specific weight to each perceptual feature to express his relative degree of importance with respect to the other features. As weights, we propose to use the the Spearman's coefficient of correlation found for each perceptual feature in our earlier work on perceptual features [1] and cited in the section on representation models. The use of such a coefficient to weight each perceptual feature allows to give more importance to perceptual features that correspond better to human judgments.

4.3.3 Hierarchical combination of perceptual features

Another original way to combine perceptual features is to consider them as a hierarchy of features. The most important features are in the highest level of the hierarchy and the less important ones are in the lower levels of the hierarchy. The features of the highest level of the hierarchy are called primary features and the features of the lower levels of the hierarchy are called secondary features. A hierarchy of several levels can be defined using the concepts of primary and secondary features. Only the features of the first level are considered as primary and have a proper existence. The features at the lower levels has no proper existence and depend on the existence of the features at the superior levels.

Taking into account these definitions and considering the definitions and estimation methods of perceptual features [1], we can deduce two hierarchical perceptual models: 1. the first model consider coarseness and directionality as primary features whereas contrast and busyness are considered as secondary features depending on coarseness; 2. The second model consider coarseness, directionality and contrast as primary features whereas busyness is considered as a secondary feature related to coarseness.

Each of these two hierarchical models can be considered in two versions: a non weighted version version and a weighted version using the same weights as in the weighted model, i.e. the Spearman's coefficient of correlation of each perceptual feature with human judgments. The weighted hierarchical model can be regarded as an hybrid model in the sense that it combines both the hierarchical and the weighted models. Finally, it must be noted that the use of a hierarchy of features creates automatically a variable weight on each of the primary features that have secondary features associated with it.

4.4 Evaluation

In this section, we take each of the similarity models proposed and compare its results to human subjects judgments of similarity. This comparison is based on the percentage of relevant images retrieved among the total number of images retrieved (relevant and non relevant) with respect to the human judgments of similarity. The methodology used can be descried as follows: 1. For each of the similarity model proposed, we observed the top 3 most similar images obtained. Then we repeat the experience successively with the 4, 5 and 6 images the most similar; 2. We consider the results of human judgments of similarity given earlier. These results are considered as judgments of relevance of responses given by each of the similarity model considered; 3. Each image, in the rankings obtained, is considered as a request. We compute the percentage of relevant retrieved images for each image request. This percentage is defined as the number of relevant images retrieved on the number of possible relevant images which can not be larger than 3 according to human subjects; 4. And finally, we consolidate the results obtained for each of the image requests into one result by summing the percentages obtained for each image request and dividing the result by the number of image requests. We obtain then the average percentage of relevance.

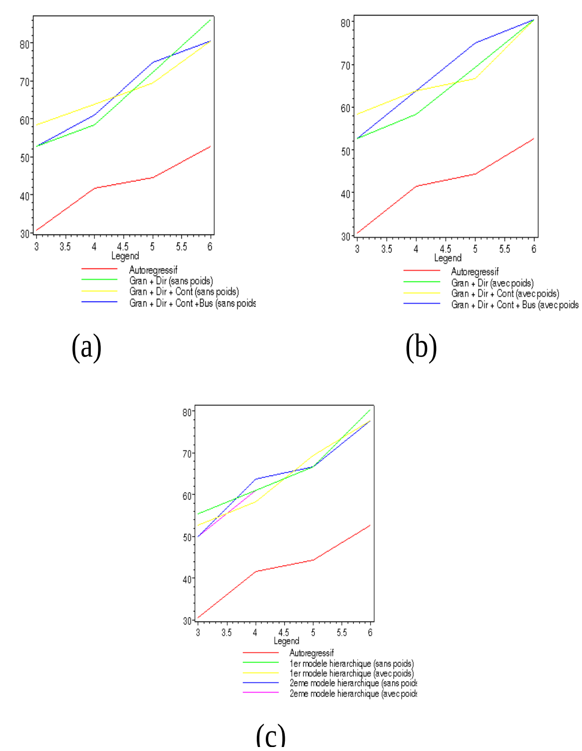

Figure 2: (a)The average percentage of relevance for the autoregressive model and the simple perceptual models. (b) The average percentage of relevance for the autoregressive model and the weighted perceptual models. (c) The average percentage of relevance for the autoregressive model and the hierarchical perceptual models.

The average percentages of relevance obtained for each of the similarity model proposed are represented in Fig. 2. A first look to these figures allow us to say that better results are obtained for all the variants of the perceptual model compared to the autoregressive model. However The results of the autoregressive model are still acceptable. Non-weighted perceptual models (Fig. 2.(a)) give almost very similar results. The addition of contrast and busyness respectively does not really change the results except the values of the similarity function become more realistic. For weighted perceptual models (Fig. 2.(b)), we found that the weighted combination of the three features, i.e. coarseness, directionality and contrast or even the weighted combination of all the four features give better results than the weighted combination of coarseness and directionality only. Results given by different hierarchical models (Fig. 2.(c)) seem to be very close to each other. However the results of the first non-weighted hierarchical model and the second weighted hierarchical model seem to be a little bit better than the two others. When the different perceptual models, non weighted, weighted and hierarchical, are compared to each other, we can say that the results are good for all of them. However The weighting of features and the organization of features in hierarchy allow to obtain more accurate results regarding the absolute values of the similarity function and the adjustment of relative positions taken by images in the similarity ranking.

5 Conclusion

A new flexible similarity model was introduced based on Gower's coefficient of similarity. The flexibility of this similarity model resides in the fact that it can be declined in several versions: non-weighted, weighted and hierarchical versions. Experimentation conducted on homogeneous textures, using both the autoregressive model and the perceptual model as content representation models, show very good results. They show clearly the superiority of the perceptual model, in all its variants, on the autoregressive model. This can be explained by the fact that, probably, perceptual similarity must rely, necessarily, on features that have a perceptual meaning; the estimated parameters of the autoregressive model have no perceptual meaning. The experimental results show also that, for perceptual similarity, the perceptual meaning of features is more important than the way to combine individual similarities to obtain a global similarity. However the way to combine individual similarities is important too. It allows mainly to obtain more realistic similarity values and also to adjust the relative positions taken by images in a similarity ranking.

References

[1] N. Abbadeni, D. Ziou and S. Wang, Computational measures corresponding to perceptual textural features, *Proceedings of the 7^{th} IEEE International Conference on Image Processing*, Vancouver, BC, September 10-13, 2000.

[2] N. Abbadeni, D. Ziou and S. Wang, Autocovariance-based Perceptual Textural Features Corresponding to Human Visual Perception, *Proceedings of the 15^{th} IAPR/IEEE International Conference on Pattern Recognition*, Barcelona, Spain, September 3-8, 2000.

[3] N. Abbadeni, *Recherche d'images basée sur leur contenu. Représentation de la texture par le modèle autorégressif*, *Research report (in French)*, N. 216, 50 p., Université de Sherbrooke, 1998.

[4] M. Amadasun and R. King, *Textural Features corresponding to textural properties*, *IEEE Transactions on Systems, Man and Cybernetics*, 19(5):1264-1274, September/October 1989.

[5] A. Del Bimbo, *Visual Information Retrieval*, *Morgan Kaufmann Publishers*, San Francisco, CA, 1999.

[6] P. Brodatz, *Textures: A Photographic Album for Artists and Designers*, Dover, New York, 1966.

[7] J. C. Gower, *A General coefficient of similarity and some of its properties*, *Biometrics Journal*, 27, pp.857-874, December 1971.

[8] R. A. M. Gregson, Psychometrics of similarity, *Academic Press Publishers*, 1975.

[9] V. N. Gudivada and V. V. Raghavan, *Content-Based Image Retrieval Systems*, *IEEE Computer*, September 1995.

[10] J. P. Guildford, *Psychometric Methods*, *McGraw-Hill*, New York, 1954.

[11] S. Santini and R. Jain, *Similarity Measures*, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 871-883, December 1999.

[12] H. Tamura, S. Mori and T. Yamawaki, *Textural Features Corresponding to Visual Perception*, *IEEE Transactions on Systems, Man and Cybernetics*, Vol. SMC-8, N. 6, June 1978.

[13] A. Tversky, Features of similarity. *Psychological Review*, 84(4):327-352, July 1977.

[14] F. Tomita and S. Tsuji, *Computer Analysis of Visual Textures*, Kluwer Academic Publishers, 1990.