

A COMPARATIVE STUDY OF TEXTURE COARSENESS MEASURES

J. Chamorro-Martínez, P. Martínez-Jiménez

Department of Computer Science and Artificial Intelligence, University of Granada
C/ Periodista Daniel Saucedo Aranda s/n, 18071 Granada, Spain
{jesus,pedromartinez}@decsai.ugr.es

ABSTRACT

There are a wide variety of measures in the literature that capture the "coarseness" texture property. Some of them have better ability to represent coarseness than the others. Furthermore, some of them are more robust against the variation of other image features, like brightness, contrast, noise and size of the image. In this paper, we propose to study the robustness and the relationship with human coarseness perception of 17 classical measures of coarseness, in order to obtain a ranking of measures. This ranking can be used to identify those measures that have the highest relationship degree with perception and the least variation with the other image features.

Index Terms— Image features, textural features, human perception, visual coarseness, fineness

1. INTRODUCTION

Texture is, together with color and shape, one of the most used feature for image analysis. It is usual for humans to describe visual textures according to some textural properties like *coarseness/fineness*, *orientation* or *regularity* [1, 2]. From all of them, the *coarseness/fineness* is the most popular one, being common to associate the presence of fineness with the presence of texture (let us remark that "coarseness" and "fineness" are opposite but related textural concepts). In this sense, a *fine* texture corresponds to small texture primitives (e.g. the image in figure 1(A)), whereas a *coarse* texture corresponds to bigger primitives (e.g. the image in figure 1(C)).

There are many measures in the literature that, given an image, capture the fineness (or coarseness) presence in the sense that the greater the value given by the measure, the coarser/finer the texture is [3]. Though most of the measures used nowadays were developed in the early nineties, they are achieving a growing interest due to applications in fields like image retrieval, where low level features like texture are the basis for the description of visual content (e.g. MPEG-7 descriptors) or for learning new concepts. In addition, new measures have been proposed during all this time.

Hence, there are several options for choosing a coarseness measure. The question is, which of them have better ability to represent coarseness, i.e., to accomplish with human coarseness perception? In particular, if we had a texture image set ordered by human subjects according to the coarseness property, we may expect a similar order if we use the computed values. Furthermore, it is desirable that the value given by a coarseness measure doesn't change with the variation of other image characteristic, like brightness, contrast, noise and size.

There are several papers in the literature where texture measures are compared [4]. However, most of these approaches do not study only coarseness measures, but a vector of measures as a whole (generally for classification purposes), and they don't take into account information about human perception of texture. One of the first papers where textural measures are analyzed taking into account human perception was presented by Tamura [2]; nevertheless, Tamura's work evaluated the degree of relationship between a single measure of coarseness and human perception. He does not provide neither a comparison between different measures of coarseness, nor an assessment about the robustness of these measures. The same limitation is found in a similar and more recent approach [5]. In other papers [6] the relationship with human perception is analyzed and different measures are compared, but they don't use the coarseness property. In [7] the authors do not identify as fundamental the coarseness property, but they focused the study in terms of image similarity, where object identification by subjects in terms of background knowledge cancel the importance of the object size (related to coarseness); in addition, as we have already explained, there are other tasks, like image description or retrieval, where coarseness has come up as a fundamental property. Moreover, in all these approaches that analyze the goodness of texture measures considering the human perception, the robustness against the image characteristic (brightness, contrast, noise and size) is not taken into account. From our knowledge, there are not studies in the literature where the robustness and the relationship with perception are analyzed together in order to compare different coarseness measures.

In this paper, we analyze a set of 17 classical coarseness measures taking into account the relationship with hu-

This work was supported by Spanish research programme Consolider Ingenio 2010: MIPRCV (CSD2007-00018).

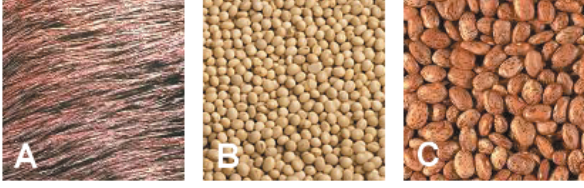


Fig. 1. Some examples of images with different degrees of coarseness

man coarseness perception as well as the robustness against brightness, contrast, noise and size. Thus, we will give a ranking, that will allow to identify those measures that have the highest relationship degree with perception and the least variation with the other image characteristics. The rest of the paper is organized as follows. The methodology used in our comparative study is described in section 2. The results and ranking are presented in section 3. Finally, the main conclusions are summarized in section 4.

2. METHODOLOGY

In this section, we introduce the methodology used in our comparative study of coarseness measures, which will be based on both (a) a study of the robustness against brightness, contrast, noise and size, (section 2.1), and (b) a study of the relationship with human coarseness perception (section 2.2). As result, a ranking of coarseness measures will be obtained.

From now on, we will note $\mathcal{P} = \{P_1, \dots, P_K\}$, with $K = 17$, the set of coarseness measures analyzed in this paper, that are shown in Table 1. This set includes classical statistical measures, frequency domain approaches, fractal dimension analysis, etc. In our study, these measures will be applied to a set of images; let $\mathcal{I} = \{I_1, \dots, I_N\}$ be the set of $N = 80$ images selected in this paper as representatives of the concept of coarseness. Figure 1 shows some images extracted from the set \mathcal{I} . The selection was done to cover the different perception degrees of coarseness with a representative number of images. Furthermore, the images have been chosen so that as far as possible, just one perception degree of coarseness is perceived¹. From now on, we will note m_k^i the value calculated by applying the measure $P_k \in \mathcal{P}$ to the image $I_i \in \mathcal{I}$.

2.1. Study of the robustness

In this section, the methodology used to study the robustness against changes in several image characteristics is described.

¹Let us remark that it is not our intention to consider 3D aspects of modeling textures [8], but the actual perception in a 2D image as would be expressed by a human user when looking at the static information present in the image. In this framework, we understand coarseness as directly related to width in number of pixels and, therefore, a zoom of an image (change in scale) will have a higher coarseness.

2.1.1. Robustness against brightness, contrast and noise

In order to study the robustness of a coarseness measure against brightness, contrast and noise, we will analyze the stability of the measure when each characteristic is progressively changed. To change a given characteristic, a parameter r , that is different in each case, will be set. In the case of brightness, r is the value added to each pixel of the image, that is, if $f(x, y)$ is the original value of the pixel, the value after change the brightness will be $g(x, y) = f(x, y) + r$. In the case of contrast, the parameter r is the multiplicative factor in the transformation defined as $g(x, y) = (f(x, y) - 128)r + 128$. Finally, in the study of noise, we will add a zero mean gaussian noise with the standard deviation as parameter.

From now on, we will note $m_k^{i,r}$ the value obtained by applying the measure $P_k \in \mathcal{P}$ to the image $I_i \in \mathcal{I}$ after changing the characteristic (brightness, contrast or noise) with the parameter r . Ideally, the value of the measure should remain almost constant when the parameter r varies slightly, that is, we expect that $m_k^{i,r} \approx m_k^i$ if r is small. But, for each measure $P_k \in \mathcal{P}$ and each image $I_i \in \mathcal{I}$, there is a critical value of r , that we will note r_c , at which $m_k^{i,r}$ starts to change dramatically. Thus, this value delimits the interval of possible values of r where the measure is stable.

In order to obtain this critical values r_c , we firstly calculate an accumulated error function defined as $AE_k^i(x) = \sum_{r=0}^x (m_k^{i,r} - m_k^i)$. Then, the Laplacian of Gaussian filter (LoG) is applied to this function, and r_c is calculated as the value where the first significant local maximum of the LoG function is reached (a maximum is considering "significant" if it is at least a quarter of the global maximum). For each measure P_k and image in \mathcal{I} , the interval $[r_s, r_c]$ will give us the set of values where the measure is stable, with r_s being the initial parameter value (corresponding to "no change"). For each measure P_k , a "range of stability" (RoS) for a given characteristic will be calculated as the mean (over \mathcal{I}) of the lengths of those intervals. We will use this "range of stability" as a goodness measure of the robustness of P_k .

In the study of brightness, we have changed the parameter r from 0 to -250 and from 0 to 250, in steps of -10 and 10, respectively. Then, two values r_c are calculated, one for positive values of r and another one for negative values. To study the contrast, the parameter r has been increased from 1 to 21.113 multiplying by 1.1, and has been decreased from 1 to 0.047 by dividing by 1.1. Thus, two values r_c are calculated (for increasing and decreasing cases). In the study of noise, the parameter r (the standard deviation of the gaussian noise) is changed from 0 to 15 in steps of 0.2. In this case, only one value of r_c is calculated.

2.1.2. Robustness against image size changes

To analyze if the values obtained for each measure are independent of the image size, windows of different sizes are

taken from the images in \mathcal{I} and the values of the measures for these sizes are compared. For each image in \mathcal{I} , a set of 100 equidistributed windows of each size are considered, using the mean of the measure values as the representative value of this size. This way, we calculate this representative value for each image with all the possible sizes between 128×128 and 64×64 . For each image, we obtain the difference of the values between all these possible sizes and we calculate the percentage of this difference with respect to the value of the original measure of this image (the one with size 128×128). If this percentage is greater than the 5% in any case, we consider that P_k is size dependent.

2.2. Study of the relationship with perception

The objective of this study is to identify the measures that best accomplish the human coarseness perception. For this purpose, we will compare the image ordering given by the measures values with the ordering given by human subjects. To obtain human assessments about coarseness perception, subjects will be asked to order the image set \mathcal{I} according to the perception degree of coarseness. Since the number of images is large and there are quite a lot of them with a similar perception degree of coarseness, we allow to order the images by assigning them to classes, where each class has associated a perception degree in increasing order. In particular, 20 subjects have participated in the poll and 9 classes have been considered.

For comparing with the order given by the pool, the values m_k^i obtained for each measure $P_k \in \mathcal{P}$ are ranked. Then, a rank statistic is used for testing the relationship between the rank given by the subjects and the rank obtained for each measure. For this purpose, we have chosen the *Spearman's rank correlation coefficient* [19], which is widely used to solve this kind of problems.

3. RESULTS

The results of the comparative study are shown in Table 1, that has been sorted in decreasing order of the ninth column. The second, third and fourth columns of this table show the "range of stability" for brightness, contrast and noise, respectively. In the case of the contrast, the values are expressed on a logarithmic scale using a base 1.1 (the multiplicative step). The fifth column shows if the measure is image size dependent or not. As we can see, four of them are affected by the size, so these measures are not taken into account in the following. The Spearman's rank correlation coefficients are shown in the sixth column. The measure of *variance* has an extremely small coefficient, so it is also rejected. The seventh column shows the "robustness score", that is obtained by combining the information of columns 2, 3 and 4. To do this, the values of the three column have been normalized by the maximum, obtaining not only a rank of measures for each

column but also a relative distance among them. Thus, the "robustness score" of each measure is obtained as the mean of the three normalized values. The eighth column shows the "perception score", that is obtained by normalizing the sixth column (Spearman's rank correlation) by the maximum. Finally, the "final score" is obtained as the mean of the "robustness score" and the "perception score" (ninth column).

From these results, we can conclude that the two best measures of this study are *Correlation* and *Tamura*, with a similar "final score". The first one has obtained best results in the study of perception, while the second one is the best in the study of robustness. The third measure in the ranking is *ED*, that has the best "perception score", but has poor results in robustness. Something similar happens with the measure of *Amadasun*, that is the fourth in the ranking. The measures of *Abbadeni* and *FD* (the fifth one and the sixth one) have an intermediate score in robustness and perception. The results obtained for the other measures are very poor in both studies.

4. CONCLUSIONS

In this paper, we have obtained a ranking of coarseness measures based on their relationship with perception and their robustness against the variation of brightness, contrast, noise and size of the image. For this purpose, 17 classical measures of coarseness have been analyzed. In order to study the relationship with perception, the values obtained by applying each measure to a set of images have been compared with the perception degree of coarseness of these images by means of a rank correlation coefficient. Pools have been used for collecting data about the human perception of fineness. In the study of robustness, we have analyzed the variation of the values of each fineness measure when the brightness, contrast, noise and size of the image are changed. From this comparative study, we conclude that six of the coarseness measures have a suitable behavior: *Correlation*, *Tamura*, *ED*, *Amadasun*, *Abbadeni* and *FD*.

5. REFERENCES

- [1] M. Amadasun and R. King, "Textural features corresponding to textural properties," *IEEE Trans. on Systems, Man Cyber.*, vol. 19, no. 5, pp. 1264–1274, 1989.
- [2] H. Tamura, S. Mori, and T. Yamawaki, "Textural features corresponding to visual perception," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 8, pp. 460–473, 1978.
- [3] R.M. Haralick, "Statistical and structural approaches to texture," *Proc. IEEE*, vol. 67, no. 5, pp. 786–804, 1979.
- [4] M. Singh and S. Singh, "Spatial texture analysis: a comparative study," *IEEE Proc. on 16th Inter. Conf. on Pattern Recognition*, vol. 1, pp. 676–679, 2002.

Fineness Measures	Brightness (RoS)	Contrast (RoS)	Noise (RoS)	Indep. with size	Perception (Spearman corr.)	Robustness score	Perception score	Final score
Tamura[2]	353.5	42.17	8.55	Yes	0.602	1.000	0.901	0.951
Correlation[3]	344.5	39.25	4.50	Yes	0.651	0.811	0.975	0.892
ED[9]	245.9	30.36	4.00	Yes	0.668	0.628	1.000	0.815
Amadasun[1]	111.8	28.81	3.07	Yes	0.648	0.453	0.970	0.710
Abbadeni[10]	212.3	16.06	4.99	Yes	0.589	0.522	0.881	0.701
FD[11, 12]	202.2	27.13	3.77	Yes	0.558	0.552	0.835	0.692
LH[3]	128.3	10.81	2.43	Yes	0.517	0.301	0.774	0.537
DGD[13]	132.1	10.26	2.01	Yes	0.509	0.284	0.762	0.521
Weszka[14]	112.1	0.00	4.61	Yes	0.482	0.285	0.721	0.506
SRE[15]	103.8	10.09	1.66	Yes	0.502	0.242	0.751	0.496
SNE[16]	92.2	12.83	2.47	Yes	0.465	0.285	0.697	0.488
Contrast[3]	99.7	0.00	4.54	Yes	0.428	0.274	0.641	0.455
Variance[3]	89.6	0.39	0.00	Yes	0.068	-	-	-
Newsam[17]	343.4	41.94	9.79	No	0.638	-	-	-
FMPS[18]	357.7	37.16	4.6	No	0.628	-	-	-
Entropy[3]	123.4	8.51	2.58	No	0.324	-	-	-
Uniformity[3]	66.0	5.42	2.70	No	0.370	-	-	-

Table 1. Ranking of measures based on the robustness of measures and its relationship with human coarseness perception

- [5] K. Fujii, S. Sugi, and Y. Ando, "Textural properties corresponding to visual perception based on the correlation mechanism in the visual system," *Psychological Research*, vol. 67, no. 3, pp. 197–208, 2003.
- [6] J.S. Payne, L. Heppelwhite, and T.J. Stonham, "Perceptually based metrics for the evaluation of textural image retrieval methods," *Int. Conf. on Multimedia Computing and Systems*, vol. 2, pp. 793–797, 1999.
- [7] A.R. Rao and G.L. Lohse, "Identifying high level features of texture perception," *Graphical Models and Image Processing*, vol. 55, no. 3, pp. 218–233, 1993.
- [8] M. Chantler and J. Wu, "Rotation invariant classification of 3d surface textures using photometric stereo and surface magnitude spectra," in *Proc. of the 11th British Machine Vision Conf.*, 2000, vol. 2, pp. 486–495.
- [9] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8, no. 6, pp. 679–698, 1986.
- [10] N. Abbadeni, N. Ziou, and D.S. Wang, "Autocovariance-based perceptual textural features corresponding to human visual perception," in *Proc. of 15th International Conference on Pattern Recognition*, 2000, vol. 3, pp. 901–904.
- [11] S. Peleg, J. Naor, R. Hartley, and D. Avnir, "Multiple resolution texture analysis and classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, , no. 4, pp. 518–523, 1984.
- [12] A.P. Pentland, "Fractal-based description of natural scenes," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 6, no. 6, pp. 661–674, 1984.
- [13] S.I. Kim, K.C. Choi, and D.S. Lee, "Texture classification using run difference matrix," in *Proc. of IEEE 1991 Ultrasonics Symposium*, December 1991, vol. 2, pp. 1097–1100.
- [14] J.S. Weszka, C.R. Dyer, and A. Rosenfeld, "A comparative study of texture measures for terrain classification," *IEEE Transactions on Systems, Man and Cybernetics*, vol. 6, pp. 269–285, 1976.
- [15] M.M. Galloway, "Texture analysis using gray level run lengths," *Computer Graphics and Image Processing*, vol. 4, pp. 172–179, 1975.
- [16] C. Sun and W.G. Wee, "Neighboring gray level dependence matrix for texture classification," *Computer Vision, Graphics and Image Processing*, vol. 23, pp. 341–352, 1983.
- [17] S.D. Newsam and C. Kammath, "Retrieval using texture features in high resolution multi-spectral satellite imagery," in *Data Mining and Knowledge Discovery: Theory, Tools, and Technology VI, SPIE Defense and Security*, April 2004.
- [18] H. Yoshida, D.D. Casalino, B. Keserci, A. Coskun, O. Ozturk, and A. Savranlar, "Wavelet-packet-based texture analysis for differentiation between benign and malignant liver tumours in ultrasound images," *Physics in Medicine and Biology*, vol. 48, pp. 3735–3753, 2003.
- [19] C. Spearman, "The proof and measurement of association between two things," *American Journal of Psychology*, vol. 15, pp. 72–101, 1904.