Texture Description based on Subtexture Components

A. Salvatella, M. Vanrell and J.J. Villanueva
Computer Vision Center - Dpt d'Informàtica
Edifici O, campus UAB
08193 Bellaterra, Spain
Email: {annasg, maria, villanueva}@cvc.uab.es

***Abstract*— The growing of multimedia content has motivated the need of tools to do image browsing and annotation; several texture descriptors have been proposed, but the high degree of complexity textures can achieve has limited their success. In this paper the concept of subtexture is introduced in order to make the automatic description of a texture adaptable to its complexity degree.**

A subtexture component is defined by sets of blobs or emergent patterns that have similar simple features, and can be fully described by a 7-dimensional vector, similar to the descriptor proposed in MPEG-7 standard. Thus, we propose a comprehensive texture description formed by the descriptions of its N subtexture components, that is, a Nx7 matrix where the number of rows is related to the complexity of the texture.

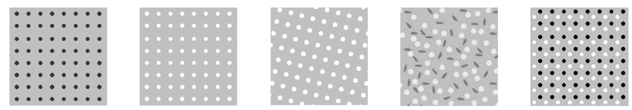
In this work a multiscale method to identify the subtexture components is presented. It is based on automatic scale selection for blob detection. Once the subtextures are identified, a global feature analysis provides the attributes of each subtexture component. Finally, the comprehensive descriptor is built from combining all subtexture information

I. Introduction

Texture is an important visual cue, and thus has been widely studied in Computer Vision, but by now a standard and general definition of texture in the Computer Vision sense has not yet been presented. Texture is necessary for many machine vision applications, and thus several computational approaches to build texture representations have been presented [1]. In most cases the representations obtained were directed by specific tasks such as image classification [2], image retrieval [3] or image segmentation [4], however psychophysical studies on human texture perception have been the motivation for others [5]. Some texture spaces have been derived from these studies, but for the moment none of the approaches leads to a general texture representation space.

Perceptual descriptions of image content are necessary to perform tasks such as browsing or annotation. A texture description in textual terms and related to how textures are perceived by human beings is necessary for image browsing or image annotation. In this scope, the MPEG-7 standard, devoted to provide a set of standardized tools to describe multimedia content, proposes a texture browsing descriptor (TBD) based on perceptual characterization of a texture [6], [7].

The goal of this paper is to present a new approach to texture description by taking into account the texture complexity degree. We try to extend the TBD descriptor by describing the subtexture components of the texture, so that all the information is comprehended and a wider and adaptable description is obtained.



(a) (b) (c) (d) (e)

Fig. 1. Examples of simple textures

To this end, the paper is organized as follows. Section 2 sets the background and section 3 defines the concept of subtexture component giving the computational details on how to obtain them. The texture description based on the subtexture components is presented in section 4. Some results are shown in section 5 and finally section 6 presents the conclusions and further work.

II. Background

As mentioned in the introduction, texture does not have a standard definition in Computer Vision. In this paper, a grey-level image is considered to be a texture if it presents homogeneity in its grey-level distribution and at least four non-overlapped windows can be taken from the image sharing the same texture properties.

Any approach to texture description should be based on how human beings perceive and describe textures. To this end, let us analyse the results that have been obtained in psychophysics on texture perception. Two approaches are confronted as being the basis for an internal visual representation of texture. On one hand, local feature extraction processes have received a hard support from the Julesz's [8] texton theory, and on the other hand, a global spatial analysis has been demonstrated to be necessary by Beck [9]. Examples in figure 1 show that both approaches are part of the process by which the human visual system deals with texture: textures in images (a) and (b) are segregable due to differences in the blob contrast, i.e. local features, whereas images (b) and (c) are segregable because of the orientation of the patterns emerging from the texture image. Therefore, not only global methods but also local properties should be taken into account when dealing with texture description.

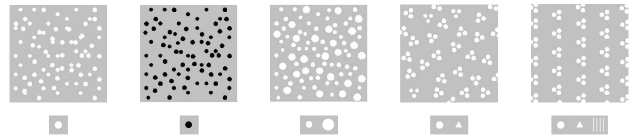
It can be shown that if textures are regarded as blobs and emergent patterns, the complexity level of textures, both natural and synthetised, is unlimited, like textures in figure 1 (d) and (e), which are made out of combination of different simpler textures, i.e. (e) is obtained by combining (a) and (b). Despite this wide range of complexity degrees in texture, in previous texture descriptors all textures are described with the same number of features. However, if human subjects are asked to describe more complex textures, they will use more words or features than they use for simpler textures.

Another advantage of considering textures as a combination of properties from blobs and emergent patterns is the ability to build objective descriptors. Most of the experiments that have been done to derive the dimensions of the texture space have been based on texture comparison or segregation. Therefore, the results that are obtained might not be suitable for texture description, but for texture comparison. Rao et al, in [10], presented a serie of psychophysical experiments concluding there are three main dimensions for texture, namely structure or regularity, scale, and directionality, nonetheless these concepts can not be clear and objective enough for description when both regular and random patterns appear in a texture at the same time. The foregoing discussion makes us consider that a texture descriptor willing to be general and meaningful should fulfil two conditions: (i) different texture degrees of complexity must be taken into account and (ii) textures have to be represented by attributes of their own characteristic elements, and not only by comparison to other textures. These considerations have motivated us to introduce the concept of subtexture component, which is defined in the following section.

III. Subtexture components

Previous considerations lead us to define a subtexture component of a texture image as a *set of blobs or emergent patterns sharing a common property all over the image*. Then, a texture image will be formed by several subtexture components, each one characterized by only one kind of blobs or emergent patterns. In figure 2 textures with different number of subtexture components are shown. The texture in image (a) has only one subtexture component defined by bright blobs randomly positioned, the image in (c) has two components due to the different size of the bright blobs and in (d) there are also two subtexture components, since there are bright blobs but also triangles emerging from the blobs grouping. Finally, texture in (e) has three subtexture components, since two previous subtexture components are positioned forming a striped emergent pattern.

The fact that textures are understood as a combination of components allows to describe textures in terms of the attributes of their components, instead of describing the whole texture. This approach to texture description fulfils the aforesaid conditions: (i) a texture can be made out of as many subtexture components as necessary, and thus the adaptation to different degrees of complexity is assured, and (ii) the subtexture components can be described in terms of the attributes of its own blobs or emergent patterns, and not by comparison with other textures.



(a) (b) (c) (d) (e)

Fig. 2. Textures having different number of subtexture components which are defined by the property presented below each image.

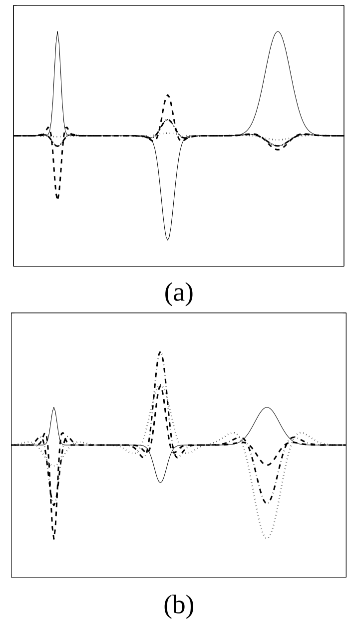
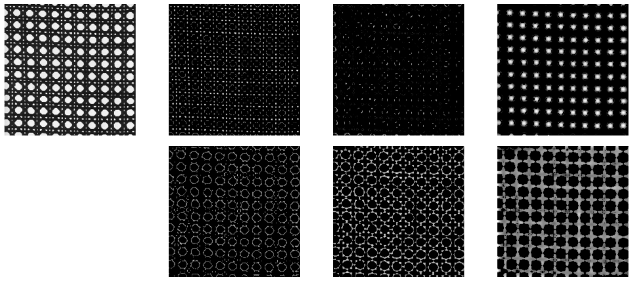


Fig. 3. Profiles of an image with 3 gaussian blobs with different contrast and σ=2,4,8 respectively. (a) shows a profile of the image (solid line) and the same profile of the image filtered by a laplacian of gaussian filter with $\sigma = 2$ (dashed line), $\sigma = 4$ (dash-dotted line) and $\sigma = 8$ (dotted line). In (b) the filter is the *normalized* laplacian operator.

Once this concept has been defined and explained, now the goal is to define a computational approach to automatically extract them since it will be the base of the texture descriptor presented in the following section.

In order to identify the subtextures forming a texture, its characteristic elements, i.e. blob or emergent patterns, have to be detected. For this purpose, the scale-space theory seems to be a good approach, since it provides a well-founded framework for dealing with image structures at different scales. Given an image, its scale-space representation is built by convolving it with gaussian kernels of different sizes [11]. Then, blobs of different sizes can be detected in this multiscale representation of the image. The problem in these multiscale approaches is to know in which scale a given blob or structure of the image is better defined.

Lindeberg in [11] proposes a method to automatically select the scale at which local image structures are better detected by differential operators. The method consists in introducing



$\sigma = 2$ $\sigma = 4$ $\sigma = 8$

Fig. 4. Extraction of subtexture components: first row shows the subtexture components obtained of bright blobs, second row for dark blobs

a new operator, the *normalized* differential operator $\partial_{\xi,norm}$:

$$\partial_{\xi,norm} = \sigma \partial_x \tag{1}$$

which corresponds to the change of variables $\xi = \frac{x}{\sigma}$. If differential operators are computed in terms of this new operator, the filter response assumes a maximum at a scale similar to the size of the local structure of the image.

For blobs detection the differential operator most commonly used is the laplacian [12], since is gives a strong response at the center of blob-like structures. In figure 3 the profiles of an image made with gaussian blobs of different sizes and contrast filtered by the laplacian and the *normalized* laplacian,

$$\nabla^2_{norm} = \sigma^2 \nabla^2 \tag{2}$$

are shown. It can be seen that when *normalized* filtering is performed the blobs are detected as maxima over space and scale, where the scale corresponds to the blob's size.

Given a texture image I, its scale-space representation is given by $\{S_\sigma(I) = I * G_\sigma\}$, where G_σ is a gaussian filter with standard deviation σ. For each scale, the filtered image $L_{\sigma_i} = \nabla^2_{norm}(S_{\sigma_i}(I))$ is computed. The bright blobs of the image can be then characterized as maxima of L_σ over space and scale, and dark blobs as minima. Therefore, given a scale, the points in L_σ which have obtained an extremal response for this particular scale are representing blobs having the same contrast and the same size. In case these blobs are homogenously distributed over the image, and according to the subtexture component definition given above, we can assume they are forming a subtexture component of the texture. Thus, if p is the number of scales considered, for an image I we obtain $n \leq 2p$ subtexture components $\{\mathcal{S}^i(I)\}_{i=1,\ldots n}$.

In figure 4 an example of how the subtexture components of a texture are isolated is shown.

IV. Texture Description

Once we have outlined the method to obtain the subtexture components of a texture, let us present the texture descriptor based on the subtexture components attributes. In [7] the TBD of a texture image is given by the regularity, two predominant directions and two predominant scales. In our case, we propose to describe a subtexture component $\mathcal{S}^i(I)$ of a texture I by

$$\mathcal{D}(\mathcal{S}^i(I)) = [c, sc, st, d_1, d_2, d_3, d_4] \tag{3}$$

where the meaning of the 7 components is the following:

- c gives the contrast of the blobs, b for bright blobs and d for dark blobs
- sc represents the scale, ranging from 1 (*small*) to 5 (*large*).
- st is the structure, ranging from 1 (*random*) to 5 (*structured*).
- d_1, d_2, d_3 and d_4 are the orientations of the predominant directions.

Let us define the steps to compute the subtexture attributes.

Contrast and scale

In previous section it has been stated that the contrast and scale of the blobs or emergent patterns forming a subtexture component are the attributes that identify it. As it has been shown, the contrast of the blobs has been derived from L_σ, and the scale is directly given by the corresponding filter.

In order to estimate the remaining features of the subtexture components we have chosen to calculate the Fourier Spectrum, which has already been used in previous works for texture feature extraction [13]. Moreover, there are psychophysical evidences that support global frequency analysis plays an important role in human perception of textures [14].

Degree of structure

In order to determine the degree of structure of a subtexture component, we will study the shape and location of its Fourier Spectrum peaks. Firstly, we will estimate a measure of the stability of them by gradually thresholding the spectrum. Afterwards, we will evaluate the alignment of the peaks by computing a modified Hough transform of the maxima, since only the lines which have been voted by several points are selected. Several measures are extracted from this analysis:

- sp : number of stable peaks (i.e. appearing in 3 or 4 thresholds)
- vsp : number of very stable peaks (i.e. appearing in 5 or more thresholds)
- l : number of straight lines

Calculation of the degree of structure,st, is then given by a weighted sum of these parameters:$\alpha \times l + \beta \times sp + \gamma \times vsp$. The values for $[\alpha, \beta, \gamma]$ have been estimated to be $[0.2, 0.3, 0.5]$ from a preliminar psychophysical experiment where 16 subjects were asked to describe textures in terms of their subtexture components features.

Predominant orientations

The predominant orientations of the subtexture components are easily detected in the spectrum, since they also appear as predominant orientations in the frequency domain. The spectrum is transformed to polar coordinates and a histogram of the orientations with 8 equally distributed bins is computed. The predominant orientations of the subtexture component are those having more than 20% of the points. This value has also been deduced from the psychophysical experiment mentioned above. The descriptor will take into account up

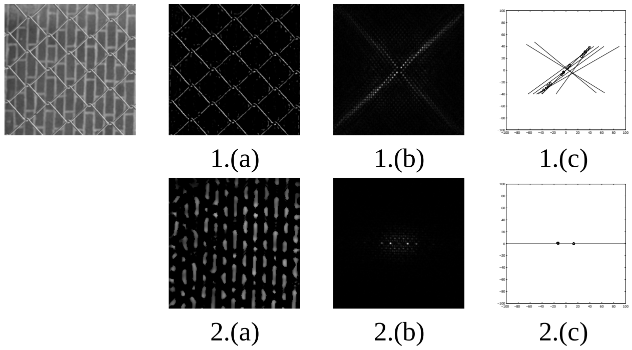


Fig. 5. Examples of subtexture components analysis for the evaluation of the degree of structure: images 1.a and 2.a are the subtexture components, their spectrums are shown in 1.b and 2.b respectively, and 1.c and 2.c illustrate the maxima and the straight lines obtained from the analysis

to 4 orientations, since it is difficult to find subtextures with more predominant directions.

Building the global texture descriptor

Since the presented computational approach can extract more than one component representing the same subtexture, we will firstly apply a selective step that removes redundant subtexture components. This redundancy is easily removed by doing a similarity test. We will denote the number of relevant subtexture components as k.

The texture global descriptor, $\mathcal{GD}(I)$ is a matrix whose rows are the description of the relevant subtextures:

$$\mathcal{GD}(I) = (\mathcal{D}(S^i(I)), \ldots, \mathcal{D}(S^k(I)))^T \tag{4}$$

As it can be seen, the number of rows of the texture descriptor depends on the texture complexity. In next section some examples of texture descriptions are given.

V. Results

The description of several textures is presented in figure 6, under every image I the corresponding global descriptor $\mathcal{GD}(I)$ is given. For example, image (a) is formed by two subtextures, one made out of bright blobs of medium scale ($sc = 3$) with an almost random structure ($st = 2$) and a predominant orientation of $135°$, and another one made out of small dark blobs with the same structure and predominant orientation. On the one hand it can be seen that the number of subtexture components that are obtained matches the complexity the texture, images (c) and (e) which can be considered complex textures are described by three and four components respectively, while images (a) and (h), which are much simpler, are described by two components only. On the other hand, we can see that the contrast, degree of structure and orientations of the subtexture components are quite well detected in most cases. Finally, it can be seen from the examples that the presented texture description is enriched by the fact that subtexture components are treated separately. For instance, in image (g) the horizontal orientation due to the emergent pattern due to elongated bright blobs is only detected for a medium scale, while the vertical orientation due to small elongated blobs appears at smaller scales.

VI. Conclusions and further work

This paper has mainly two contributions. Firstly, the concept of subtexture component has been introduced, which allows a texture description that can be interesting both from a computational and a perceptual point of view. Secondly, we have presented a first approach to a computational texture descriptor which is shown to be general enough to give the description of any natural texture.

The fact that the number of subtexture components can vary makes this approach suitable to all levels of texture complexity, which is very important for Computer Vision applications where all types of images can be found. The presented texture descriptor is based on perceivable characteristics of the image without the need of comparison. This is indispensable for applications such as image browsing where images have to be described in terms of its own properties and in a way that makes it easy to go from natural language to computational representations.

Further work will be focused on the introduction of more complex information such as the shape of the emergent patterns and on an improvement of the subtexture components descriptor. Besides, the performance of the description given in natural language terms for image browsing in real applications still has to be tested.

References

[1] Tuceryan, M., Jain, A.K.: *Handbook of Pattern Recognition and Computer Vision*, chapter Texture Analysis, pages 235–276. World Scientific, 1993.

[2] Ojala, T., Pietikainen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. *PR*, 29(1):51–59, 1996.

[3] Manjunath, B.S., Ma, W.Y.: Texture features for browsing and retrieval of image data. *IEEE-PAMI*, 18(8), 1996.

[4] Boyer,K., Sarkar, S. editors: *Perceptual organitzation for artificial vision systems*, chapter 9, pages 139–172. Kluwer Academic Publishers, 2000.

[5] Tamura,H., Mori, S., Yamawaki,T.: Textural Features Corresponding to Visual Perception *IEEE Trans. Syst. Man Cybern*, 8:460–473, 1978.

[6] Wu,P. ,Manjunath, B.S., Newsam,S. Shin, H.D.: A texture descriptor for browsing and similarity retrieval. *Journal of Signal Processing: Image Communication*, 16:33–43, 2000.

[7] Manjunath, Salembier, and Sikora, editors. *Introduction to MPEG-7.Multimedia Content Description Interface* . John Wiley and Sons, 2002.

[8] Julesz, B., Bergen, J.R.: Textons, the fundamental elements in preattentive vision and perception of textures. *Bell Systems Techn. Journal*, 62:1619–1645, 1983.

[9] Beck, J.,Sutter, A.Ivry, R.: Spatial frequency channels and perceptual grouping in texture segregation. *CVGIP*, 37:299–325, 1987.

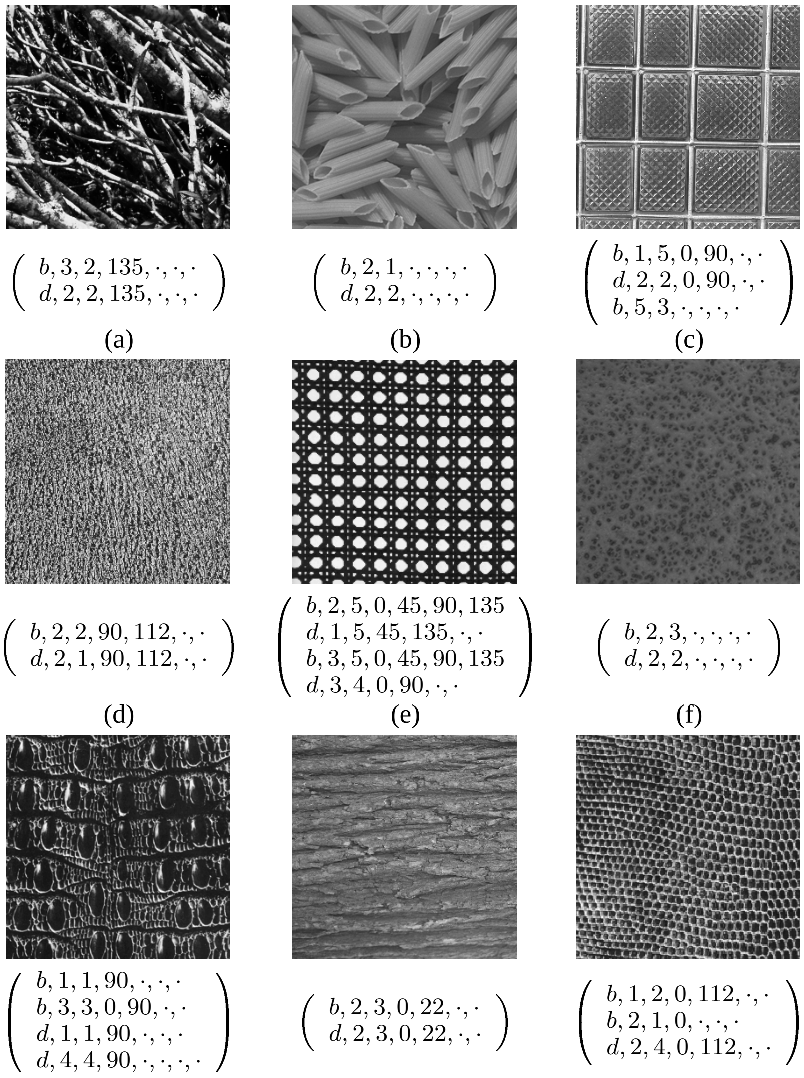
[10] Rao,A.R. ,Lohse, G.L.: Towards a texture naming system: Identifying relevant dimensions of texture. *Vision Research*, 36:1649–1669, 1996.

[11] Lindeberg, T.: Scale-Space Theory in Computer Vision. Luwer Academic Publishers, 1994.

[12] Voorhees, H., Poggio, T.: Detecting Textons and Texture Boundaries in Natural Images, International Conference on Computer Vision, 1987, 250-258.

[13] Liu, F., Picard R.W.: Periodicity, directionality, and randomness: Wold features for image modelling and retrieval. *IEEE Trans. on PAMI*, 18(7):722–733, 1996.

[14] Harvey, L.,Gervais, M.: Visual texture perception and Fourier analysis. *Perception and Psychophysics*, 24(6):534–542, 1978.



$$\begin{pmatrix} b, 3, 2, 135, \cdot, \cdot, \cdot \\ d, 2, 2, 135, \cdot, \cdot, \cdot \end{pmatrix}$$

(a)

$$\begin{pmatrix} b, 2, 1, \cdot, \cdot, \cdot, \cdot \\ d, 2, 2, \cdot, \cdot, \cdot, \cdot \end{pmatrix}$$

(b)

$$\begin{pmatrix} b, 1, 5, 0, 90, \cdot, \cdot \\ d, 2, 2, 0, 90, \cdot, \cdot \\ b, 5, 3, \cdot, \cdot, \cdot, \cdot \end{pmatrix}$$

(c)

$$\begin{pmatrix} b, 2, 2, 90, 112, \cdot, \cdot \\ d, 2, 1, 90, 112, \cdot, \cdot \end{pmatrix}$$

(d)

$$\begin{pmatrix} b, 2, 5, 0, 45, 90, 135 \\ d, 1, 5, 45, 135, \cdot, \cdot \\ b, 3, 5, 0, 45, 90, 135 \\ d, 3, 4, 0, 90, \cdot, \cdot \end{pmatrix}$$

(e)

$$\begin{pmatrix} b, 2, 3, \cdot, \cdot, \cdot, \cdot \\ d, 2, 2, \cdot, \cdot, \cdot, \cdot \end{pmatrix}$$

(f)

$$\begin{pmatrix} b, 1, 1, 90, \cdot, \cdot, \cdot \\ b, 3, 3, 0, 90, \cdot, \cdot \\ d, 1, 1, 90, \cdot, \cdot, \cdot \\ d, 4, 4, 90, \cdot, \cdot, \cdot, \cdot \end{pmatrix}$$

(g)

$$\begin{pmatrix} b, 2, 3, 0, 22, \cdot, \cdot \\ d, 2, 3, 0, 22, \cdot, \cdot \end{pmatrix}$$

(h)

$$\begin{pmatrix} b, 1, 2, 0, 112, \cdot, \cdot \\ b, 2, 1, 0, \cdot, \cdot, \cdot \\ d, 2, 4, 0, 112, \cdot, \cdot \end{pmatrix}$$

(i)

Fig. 6. Examples of texture descriptions