CVC Tech.Rep. #60 January, 2002

Towards a Texture Representation Database

Anna Salvatella and Maria Vanrell

Computer Vision Center / Dept. Informatica (UAB)
Edifici O - Campus UAB
08193 Bellaterra (Barcelona)
annasg,maria@cvc.uab.es

Submitted to the Computer Vision Center
on January, 2002

Contents

Towards a Texture Representation Database

by

Anna Salvatella and Maria Vanrell

Abstract

In this paper we propose a texture description method to build a wide representation dataset that allows validating the performance of different computational texture representations. Some psychophysical data has been collected in previous works, where a set of textures have been observed and described by a set of subjects, all this experiments have been done in closed labs and on different sets of image textures. Due to the lack of a common accepted texture representation, most of the experiments have been based on similarity judgements between textures or on quantification of high-level features. The former approach is usually closed to a restricted set of images and the latter is subject to certain degree of subjectivity and metamerism. In this paper we propose a blob-based representation that intends to put the basis to do open psychophysical experiments on texture images with no constraints on the set of selected images neither on the dependence on high level linguistic terms. Following a blob-based approach we demonstrate the possibility to derive computational representations to be easily tested with the collected data.

1 Introduction

Most of the papers on texture perception and representation refer to the lack of a standard representation space. In computer vision a great deal of different representations have been developed. They are based on different mathematical approaches trying to define a general texture representation [8, 13, 39] where the approaches based on Gabor filtering seems to become the most successful in terms of generality [18, 26]. In psychophysics much work has been focused on trying to understand how the perception of texture occurs from different points of views, preattentively [19, 22, 2], and attentively [14, 30, 16, 17].

This problem is emphasized when texture is compared to colour, since both share the fact of being surface properties. While the latter is physically modelled, psychophysically measured and completely tabulated on standard spaces that can be appointed by anyone [38], the former is lacking of all of them. Colour and texture of a surface are physically provided by two factors, the spectral power distribution of the incident light and the spectral reflectance of the surface. The position and light sensitivity of the observer add some important properties.

In this paper we will only refer to the problem of representation of image texture, we will consider as different textures those images that preattentively could be discriminated by human perception. We will regard as different textures those images corresponding to the same surface where any of the above parameters, which contribute on image formation, have been changed. A relevant work is being done in characterising reflectance of texture surfaces [9].

This work is a proposal of a texture representation system based on the objective concepts of blob and emergent pattern. It provides us with two basic advantages. Firstly, to give the opportunity to build a consistent dataset of psychophysical data on how texture is perceived by a large number of different individuals, since it can provide a large perceptual space for which we can be able to find a 2nd order isomorphism in the Edelman's sense [10], it is commonly done in psychophysical works [14, 16] where perceptual spaces are correlated with computational spaces [36]. Secondly, to establish a guideline on how computational texture representations should be built in order to directly match with all the collected data.

To this end, in this paper we propose an experiment to be realised on a web site[1] devoted to collect all the data about the questions we propose in this paper on how an image texture is perceived. We propose to do the experiment using a web page because it will allow to collect a large amount of experiments on different subjects and on wide sets of images since it can remain open for a long time and available from anywhere.

The paper has been organized as follows. Firstly we give an overview on previous experimentation done in psychophysics related to texture representation problem. Secondly, we give a brief review on how the evaluation of texture representation has been done in computer vision. Thirdly, we give some definitions as the basis to introduce our proposal of a representation based on the terms of blob and emergent pattern. In the next section we explain how the experiment will be posed and controlled. Finally, we suggest a computational representation to demonstrate that is feasible to automatically extract all the data collected with the proposed experiment. As a final remark we want to note that all this work has to be understood in the frame of a thesis project in its beginning and from here our interest in making evident its proposal character, since a lot of details still remain to be closed.

2 Previous works

In computer vision, the effort on texture research has been mainly directed to develop mathematical approaches and algorithms that decompose and measure texture properties in order to perform specific tasks of classification or segmentation [8, 13, 39]. Some other works have focused on defining computational models pursuing a behaviour that agrees with physiological mechanisms of early vision [25, 4, 32] and match psychophysical data.

[1]This web site is still under construction, please contact to correspondence author to get the final address.

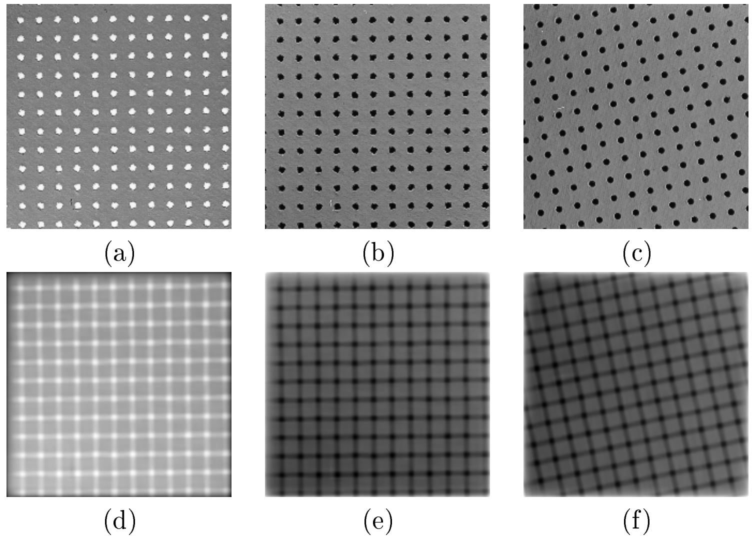


Figure 1: Examples of textures formed by simple blobs and their emergent patterns.

In psychophysics, there are two main ways to behave when the texture perception problem is studied, an excellent review work can be found in [3]. The difference between them is derived from the experimental techniques used to get the empirical data, whereas some works are based on the phenomenon of preattentive segregation of two overlapped textures [20, 1] , others are based on the analysis of similarity judgements between pairs of textures [14, 31, 16, 17] that is based on an attentive process.

Results from both types of approaches are conditioned by many factors that make arise some questions. Firstly, how the selected set or pairs of images used can modify the performance measured by the experiment?, similar experiments on colour naming are exclusively addressed to analyse the influence on the results from the set of selected stimulus [6, 33]. Secondly, how the a priori knowledge on the image scene of the subjects can affect similarity judgements between natural images used in some experiments? as it is done in [31, 34]. A recent work on natural images [17] concludes with the context-dependency of the experiments based on an attentive similarity judgement.

To avoid the last problem some authors have based their experiments on synthetic images [14, 15]. From a reasonable point of view it seems quite evident that it is impossible to do an algorithm able to generate any possible texture image. If it was possible, it would allow to build the general set of texture images as it has been done in colour research, where it is easy to synthesise all the coloured lights within the visible spectrum interval.

Despite the limitations introduced by the experimental techniques, important results have arisen from psychophysical research on texture perception, where two different approaches are confronted as being the basis of a visual internal representation of stimuli. On one hand, the feature extraction process has received a hard support from the Julesz's texton theory [20, 21], that concludes texture discrimination is due to differences on first order statistics of textons, which have been defined as the image blobs and their attributes. On the other hand, a global spatial frequency analysis seems to be indispensable to be able to capture the segregation of textures due to patterns emerging from the arrangement of the image blobs, as it is argued in [2, 16]. In figure 1 we can see both cases, that is, texture differences due to differences on blob attributes, as is the contrast in images (a) and (b), and differences due to differences on attributes of emergent patterns, as is the orientation in images (b) and (c).

In the present work we clearly advocate for a combination of both processes as the best way to understand texture segregation and representation. It is also done in most of the computational works where representation is computed by convolving the image by a bank of frequency-selective filters [18, 25, 5]. This multi-scale filtering allows performing both processes, that is, tuning in image blobs or in emergent patterns depending on the scale of the filter used.

3 Texture representation evaluation

As we have previously introduced, a common approach in computer vision is to give a texture representation based on the energy computation of the outputs of a bank of filters. Following this approach or any other else where the output is a numerical vector representing texture properties, the evaluation has been usually done by computing an index of retrieval performance [26, 12] or a classification error rate [28, 29].

The accuracy on texture image retrieval is usually evaluated on standard texture database as the VisTex [35] or the Brodatz's album [7]. Homogeneous textured images are divided into N small subimages, creating a large image set. A perfect retrieval is done when the similarity measurement between the entire image database gives the N higher values for those subimages corresponding to the same query image. Then, a usual performance measurement is given by the average number of subimages belonging to the same query image in the first N most similar. A classification error rate can be computed in a similar way when a classifier is used to associate each subimage to an image class. In both cases, validation informs about the performance of a representation in finding equal textures, but it does not inform on how good it is in giving automatic similarity judgements.

A less common method to validate a computational texture representation has been to identify relevant dimensions on a low-dimensional texture space, for which a meaningful interpretation can be given from a perceptual point of view. Therefore, a computational representation is derived by joining statistical measurements for each dimension. Although it was done in last seventies, the work presented by Tamura et al. in [34] is still current from a methodological point of view and of great value from the performance evaluation issue. They selected a subset of statistical measurements on textures properties based on good correlation with judgements done by human participants in a experiment. It is based on judgements of similarity of some individuals and as we already mentioned before, they suffer from the context-dependency of the experiments based on an attentive perception process and from the constraint imposed by a closed set of images.

Other works, have validated the set of selected features by directly using those which correlates with high-level features associated to a experimentally deduced perceptual space, as is the case of Picard et al in [11], where measurements are selected following the dimensions identified by Rao et al. in [31].

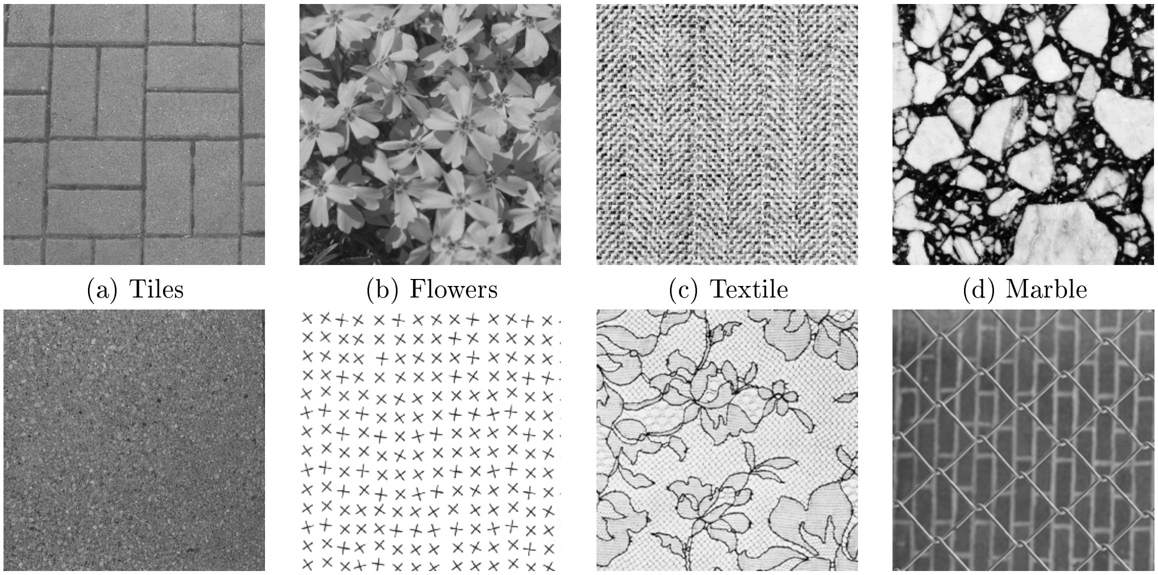
Finally, in a more recent work by Manjunath et al [40], a validation is done on five individuals who are asked to quantify five aspects of a texture, one for texture structure, two for predominant directionalities and two for predominant scales. This data is correlated with the data of their algorithm that provides a five-dimensional texture representation derived from the analysis of the output of a Gabor filter bank on the input image. From our point of view, the most important contributions of this work are two, firstly, they propose a low-dimensional space where dimensions are associated to quite objective image properties, and secondly, they propose psychophysical experimentation for each dimension.

The above two statements are the starting point for the proposal of a psychophysical experiment based on answering concrete questions about objective properties of the image. We seek to insert objective properties by building a texture definition based on image blobs and emergent patterns. Before doing the experiment, individuals will have to familiarize themselves with these two concepts.

With our proposal we try to avoid some metamers[2], that could be derived from a five-dimensional representation as the one previously presented, we assume as first priority to seek for a representation that allows to capture any small difference between any pair of textures.

For instance, a texture with four predominant directions could become a metamer if we only consider two dimensions to represent directionality, as the image in figure 2.(h), a similar case can occur if we

[2]We have adopted a term from colour science, metamer, to refer to perceptually different image textures but sharing a common representation.



(a) Tiles (b) Flowers (c) Textile (d) Marble

(e) Cork (f) Crossings (g) Lace (h) Bricks and fence

Figure 2: Texture images

only consider two predominant scales, we can immediately find or synthetically build a new texture with more than two important scales, as it is shown in figure 2.(d). We propose to solve metameric problems by introducing a representation based on the concept of subtextures.

4 Blob-based texture representation approach

The goal of this section is to give a methodology to build texture representations based on the image blobs. Before giving some basic concepts we want to remark the purpose of combining the two main conclusions from psychophysical research on texture perception which are summarized in the Julesz's texton theory [22] and by the evident need of global spatial-frequency analysis to deal with emergent patterns [2].

Blobs are the basic elements of early visual representations. Although they were earlier defined in computer vision in the Marr's primal sketch [27], subsequently they have been redefined by Lindeberg in [24]. Since the mathematical definition is too complex to be explained to non-expert people willing to do an experiment, we need to give a plain definition of the concept of *blob* as a *closed image region whose pixels share the properties of grey level homogeneity, compactness and convexity*, examples of image blobs are shown in figure 4.

Different attributes can be associated to image blobs, blobs can be *dark* or *bright* in reference to the predominant image grey level or blob neighbourhood, blob *contrast* is also an important attribute that is given by the relationship between the grey levels of the blob and its neighbours. An important attribute of image blobs is its *scale* or size that is directly related to the blob area. Some blobs will also present a wide difference between their own principal axes; these will be elongated blobs or *bars* for which an *orientation* attribute will be associated, this will be the orientation of the first principal axis. Blobs, bars and their attributes correlates with an important part of textons introduced by the texton theory.

Next, we need to introduce the image *background* that is assumed to be those parts of the image that can not be perceptually considered as independent blobs due to they are perceived as *a wide connected region across the image.*

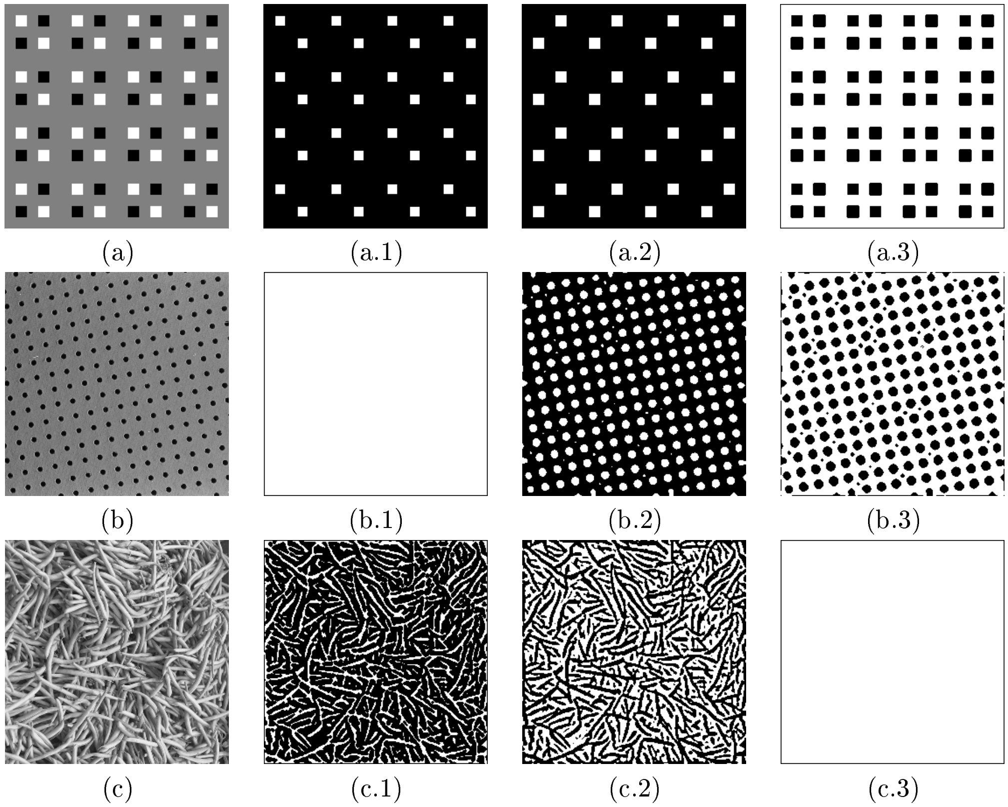


Figure 3: (a) Synthetic image. (b), (c) Natural images. (x.1) Bright blobs of image x. (x.2) Dark blobs of image x. (x.3) Background of image x.

In figure 3 we can see the decomposition of three images in terms of their bright blobs, dark blobs and background. An ideal case is shown in image 3.(a), formed by perfect bright and dark blobs. Image 3.(b) does not present any bright blob, and image 3.(c) does not have background. In this last case we can clearly see that blobs represent homogeneous image regions that are meaningless entities for an attentive perceptual process.

However, blobs and background are not enough properties to represent texture, as we have already shown with images (b) and (c) of figure 1, where these two different textures present identical blobs and therefore identical 1st order statistics in the sense of texton theory. Difference arises from an emergent pattern perceived from the blob arrangement, hence, a global spatial-frequency analysis is needed [2]. At figure 1. (e) and (f) we display the checked pattern emerging from both textures.

From the previous example we can conclude that spatial blob organisations can be essential to represent textures and to avoid metamers. Therefore, we need to give a definition for emergent patterns, since our representation will have to consider their existence and attributes. We can assume that an *emergent pattern* is an *association of blobs that are perceptually grouped as entities that appear repetitively across the image.*

As we have done for blob attributes, we will also consider the attributes of the emergent patterns. We will describe every emergent pattern in the image by its *orientation*, *scale* related to the image size and a degree of *confidence* about how it is perceived. We will give a low degree of confidence when they are only intuit, as the checked pattern emerging from image 1.(a),(b). High degrees of confidence will be assigned when the checked pattern is perfectly defined as in figure 2.(a).

At this point, difficulty arises when we need to define which are the most common emergent patterns that can be perceived from any natural or synthetic image. Some research has been done in this sense and some more will need to be done in future, but the experiment we propose here can be a good test-bed to analyse this problem. As a starting point, we propose the following emergent patterns: *ring, striped, checked, crossing, arrow, T* and *polygonal*. In figure 6 we can see the proposed emergent patterns as a basis to describe any texture. In section 6 we will show how the automatic detection of emergent patterns will require scale-space mechanisms, as it can be seen in figure 8.

Once we have defined blobs and emergent patterns we need to compile all the given information. We propose to do it by introducing the concept of subtexture. A *subtexture* is *a set of blobs or patterns sharing a common attribute*, such as scale, orientation and contrast or confidence. A global texture representation will be obtained by joining all its subtexture representations. The number of subtextures of a given texture can vary depending on its complexity.

To show the ability of subtexture concept in describing textures, let us comment some details on how different images in figure 2 will be represented later on. The results are summarized on tables 5, 6 and 7. An example on how the number of subtextures explains the texture complexity is shown with image (f), only two subtextures are needed to represent it, one to describe elongated dark blobs and the other one to explain the crossing emergent pattern, on the other hand five subtextures are needed to describe images (c), (d) and (g).

Image (d) is described by four subtextures, three of them to describe three different sizes of bright blobs, and one subtexture to describe their polygonal emergent shape. Apart from subtextures description, images (d), (e) and (f) present the property of having a background. Images (a) and (h), share the property of having T emergent patterns that usually appear like T-junctions in brick walls. The confidence of these emergent patterns is very high in both images, whereas the ring pattern in image (b), formed by the petals of the flowers, has a very low confidence.

5 Proposed experiment

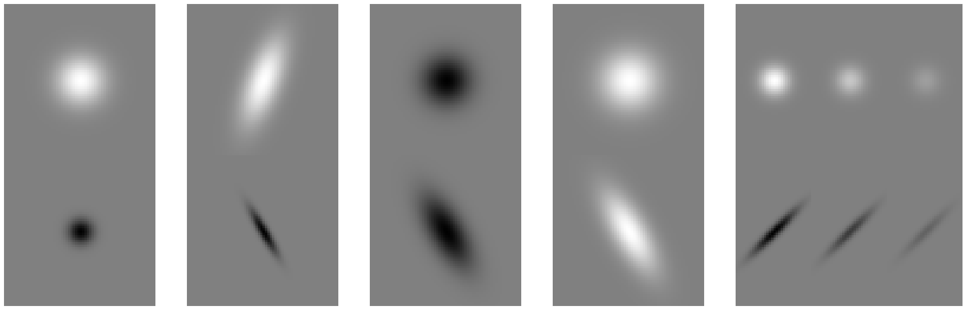
According to the blob-based representation we have defined above, we propose now a psychophysical experiment to collect information about human texture perception following a subtexture scheme in terms of their blobs and emergent patterns.

The first step to assure consistency of data concerns fixing display conditions of any texture. As we are suggesting a web-based experiment, that is an open laboratory, every experiment will have to pass a calibration process to assure equal conditions on any screen. Because our image database can change along time, we will not be able to assure a fixed image size, therefore we have selected a visual angle that allows to see a usual 256×256 image on a common display resolution. The selected conditions we propose are the following, the image being the stimulus should be displayed on 7.5 degrees of visual angle and from a viewing distance of 50cm. The stimulus should be overlapped on a 15 degrees background with a grey level equal to the mean of the stimulus displayed, this is suggested to avoid influences from the background, although a common background for all the images could also be justified in order to fix contrast judgements. The procedure for the calibration process can be easily implemented by printing a standard pattern that has to be manually suited to the window size whose final conditions can be stored by the web software in Java.

The experiment can not begin unless the user has performed two actions, these are, the calibration process and reading the necessary a priori knowledge on how to perform the experiment, that is, how to look at texture. Thus, concepts of blobs, emergent patterns, attributes and subtextures have to be understood, it can be done by selecting a subset of suited examples, for which a blob-based representation is explained. Previous experiences have been done in closed laboratories and the experience has been successful with non expert subjects.

Once all the previous steps have been done, subjects are asked to answer the following questions:

1. How many different types of blobs do you see in this texture?



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

Figure 4: (a) Blobs. (b) Elongated blobs or bars. (c) Dark blobs. (d) Bright blobs. (e) Different contrast blobs, High, Medium and Low, from left to right.

Blob Shape		*Blob Contrast*					*Blob Scale*				
Non-elongated	*Elongated*	*Dark*	*Bright*	*High*	*Medium*	*Low*	*Very Big*	*Big*	*Normal*	*Small*	*Very small*
○	○	○	○	○	○	○	○	○	○	○	○

Table 1: Multiple-choice answer form to fill in after question 1.

2. Do you perceive an homogeneous background apart from the blobs?
3. How many different types of emergent patterns do you perceive in this texture?
4. Are the elongated blobs and emergent patterns sharing one or several common orientation?
5. How is the organisation of blobs and emergent patterns across the image?

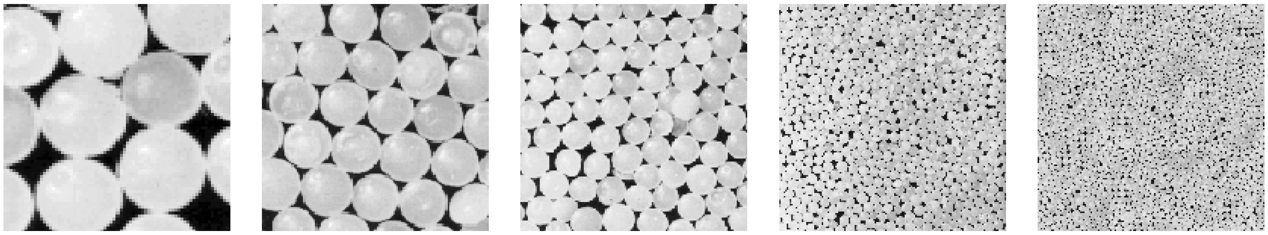
For every question some other sub-questions will arise, all of them require a multiple-choice answer that is collected by a form. Let us examine the information collected by each question.

The aim of questions 1 and 3 is to get all the information from different image subtextures. In question 1 subjects have to answer about subtextures formed by a specific type of blobs. They have to describe their shape, that is to decide if blobs are elongated or not. They also have to describe contrast, it implies to point out if blobs are dark or bright, and if they present high, medium or low contrast. All these attributes are illustrated in figure 4. Finally they have to describe scale, the election of the scale will be made by similarity to other images where scale is the essential texture attribute. These images are shown in figure 5. A less intuitive way to specify scale could be by a relative scale factor between the size of the image and the image window containing the blob. From our experience, the comparison between differently scaled images is a quite easy way to answer the question.

The goal of question 3 is to collect information about subtextures due to emergent patterns. Subjects have to answer about the shape of the pattern chosing on the corresponding shape, as it is shown in table 2 and considering theoretical emergent patterns shown in figure 6.

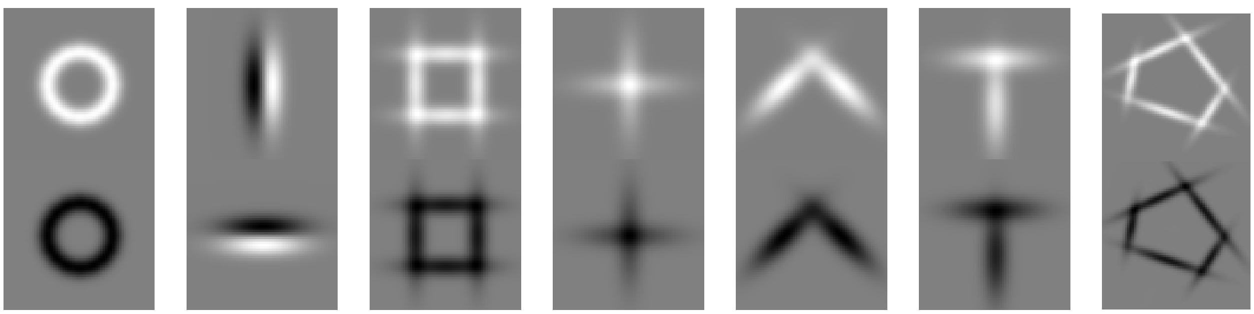
For each type of pattern subjects are asked to give a degree of confidence on how this pattern is perceived. The answer should be 5 if the subtexture is formed by perfect rings, however it should be 1 if the subtexture is formed by 5 blobs perceptually grouped on a circumference. As it has been done for blob-based subtextures the scale of patterns has to be specified. All this information is compiled by the form presented in table 2.

Above questions and corresponding forms are those from which a major difficulty raise and probably greater variances will arise from judgements. The rest of questions are easier to answer. Question 2 gets information about the existence of a background, figures 3.(b) and (c) show two examples of textures with and without a perceived background, respectively. Question 4 asks for the existence of predominant



(a) Very Big (b) Big (c) Medium (d) Small (e) Very small

Figure 5: From left to right, examples of textures with very big scale, big scale, medium scale, small scale and very small scale.



(a) (b) (c) (d) (e) (f) (g)

Figure 6: (a) Ring. (b) Striped. (c) Checked. (d) Crossing. (e) Arrow. (f) T. (g) Polygonal.

Pattern Shape							*Pattern Confidence*					*Pattern Scale*				
Ring	*Striped*	*Checked*	*Crossing*	*Arrow*	*T*	*Poly.*	*1*	*2*	*3*	*4*	*5*	*Very Big*	*Big*	*Normal*	*Small*	*Very small*
○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○	○

Table 2: Multiple-choice answer form to fill after question 3.

0°	22°	45°	67°	90°	112°	135°	157°	*Random Orientation*
○	○	○	○	○	○	○	○	○

Table 3: Mulitple-choice answer form to fill after question 4.

Perfectly ordered	*Almost ordered*	*Certain order*	*Almost random*	*Completely random*
○	○	○	○	○

Table 4: Mulitple-choice answer form to fill after question 5.

directions in subtextures, which can be described by sticking on the form presented in table 3. Most common oriented subtextures usually present one or two predominant directions, although if there are more they can be collected in the proposed form. In case texture present random orientation, it can also be specified in the form.

Last question tries to describe the structure of a subtexture. We have selected five degrees of structure for each subtexture (See figure 7).The form for this question is specified in table 4, we suggest a possible scale for texture structure in figure 7. Two criteria can be used to establish the degree of structure: the fulfilment of perfect rules for arranging patterns across the image and the existence of a perfect pattern.

In tables 5, 6 and 7 we give some examples on how a texture can be completely specified by answering all the questions of the proposed experiment.

6 Feasibility of a computational representation

In this paper we have given a guideline on how to collect some objective data for texture description. Now, we want to demonstrate the possibility to derive computational representations to be easily tested with the collected data.

Since the proposed representation is built from a subtexture interpretation, we propose a three steps algorithm that perform the following processes:

Subtexture isolation process based on a multi-scale laplacian of gaussian filtering.

Global spatial-frequency analysis from the Fourier spectrum of the filtered images.

Interpretation of subtexture properties based on the Hough transform of the previously computed information.

To deal with the first process, we need to decompose the image in three parts, as we want to consider bright blobs, dark blobs and background separately. The method used to make this decomposition is the laplacian of gaussian filtering.

In a first step to obtain a complete description of an image texture, I, it is convolved by a gaussian function in order to tune the filter scale with the size of the different blobs in the image, that is

$$I' = I * \nabla^2 G_{\sigma_i} \;\; where \;\; \sigma_i \in [0.5, 3], i = 1...5 \tag{1}$$

and

$$G_\sigma(x,y) = \frac{1}{2\pi\sigma^2} e^{\frac{x^2+y^2}{2\sigma^2}} \tag{2}$$

We are assuming a computational definition of blob similar to the one given in [37], where they are defined as the duals of edges. Therefore, blobs are detected by finding the zero crossings of the filtered image. Two main reasons justify this method, the first one is that the zero crossings assure closed blob

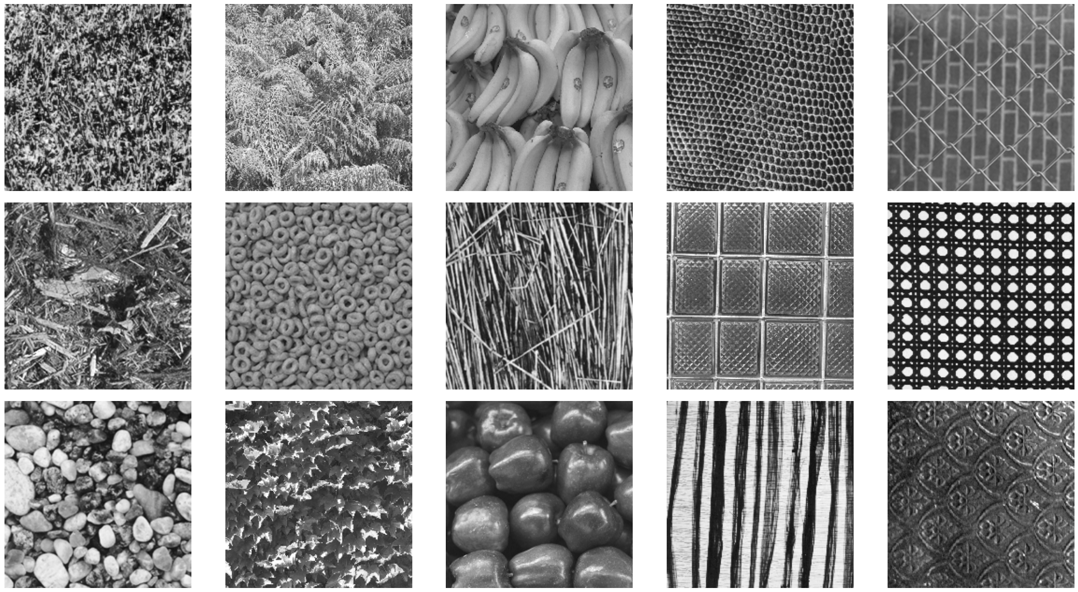


Figure 7: Examples of images with different degree of structure. Image columns are organized from left to right from random to perfectly ordered.

Image	*Blob Shape*		*Blob Contrast*					*Blob Scale*				
Subtexture	*Non-elong.*	*Elong.*	*Dark*	*Bright*	*High*	*Medium*	*Low*	*V.Big*	*Big*	*Medium*	*Small*	*V.Small*
Tiles.B.1		√	√			√		√				
Tiles.B.2		√		√		√		√				
Flowers.B.1		√		√		√				√		
Flowers.B.2	√		√			√	√				√	
Flowers.B.3	√		√		√	√				√	√	
Textile.B.1		√	√			√				√		
Textile.B.2		√		√		√				√		
Marble.B.1	√			√	√						√	
Marble.B.2	√			√	√					√		
Marble.B.3	√			√	√				√			
Cork.B.1	√		√				√					√
Cork.B.2	√			√			√					√
Crossings.B.1		√	√		√						√	
Lace.B.1		√	√			√					√	√
Lace.B.2		√	√		√					√		
Lace.B.3	√			√			√			√		
Brick&Fence.B.1		√	√			√				√		
Brick&Fence.B.2		√		√		√				√		
Brick&Fence.B.3		√		√	√				√			

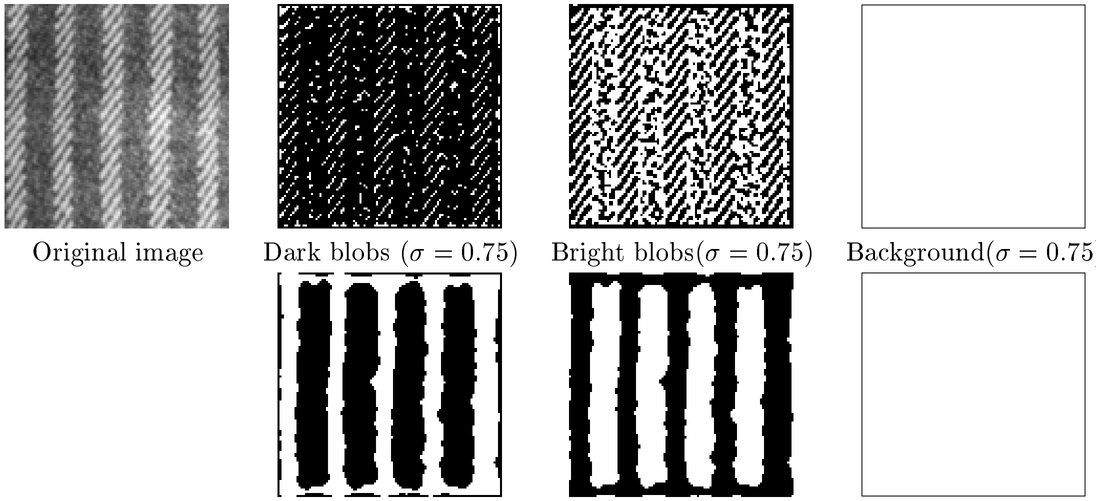
Table 5: Examples of blob-based subtexture representation for texture images in figure 2.

Image	Pattern shape							Pattern Confidence					Pattern Scale				
Subtexture	R.	S.	Ch.	Cr.	A.	T.	P.	1	2	3	4	5	V.B.	B.	M.	S.	V.S.
Tiles.P.1						√						√	√				
Tiles.P.2				√								√	√				
Tiles.P.3							√					√		√			
Flowers.P.1	√								√					√			
Flowers.P.2	√									√						√	
Textile.P.1		√										√			√		
Textile.P.2		√								√			√				
Textile.P.3					√			√					√	√			
Marble.P.1							√				√			√			
Marble.P.2							√			√					√		
Crossings.P.1				√								√				√	
Lace.P.1			√								√					√	√
Brick&Fence.P.1						√						√			√		
Brick&Fence.P.2			√									√		√			

Table 6: Examples of pattern-based subtexture representation for texture images in figure 2. The pattern shapes are: Ring (R.), Stripped (S.), Checked (Ch.), Crossing (Cr.), Arrow (A.), T (T.) and Polygonal (P.). The scales are: Very Big (V.B.), Big (B.), Medium (M.), Small (S.) and Very Small (V.S.).

	Orientation								Structure				
Image Subtexture	0°	22°	45°	67°	90°	112°	135°	157°	Random Orient.	P. O.	A. O.	C.O.	A.R.	C.R.
Tiles.B.1	√				√					√				
Tiles.B.2	√				√					√				
Tiles.P.1	√				√					√				
Tiles.P.2	√									√				
Tiles.P.3	√				√					√				
Flowers.B.1									√				√	√
Flowers.B.2									√			√	√	
Flowers.B.3									√					√
Flowers.P.1									√					√
Flowers.P.2									√					√
Textile.B.1			√				√				√			
Textile.B.2			√				√				√			
Textile.P.1			√				√			√				
Textile.P.2					√					√				
Textile.P.3	√									√				
Marble.B.1									√					√
Marble.B.2									√					√
Marble.B.3									√					√
Marble.P.1									√					√
Marble.P.2									√					√
Cork.B.1									√					√
Cork.B.2									√					√
Crossings.B.1									√			√		
Crossings.P.1									√			√		
Lace.B.1			√				√				√			
Lace.B.2									√					√
Lace.B.3									√				√	
Lace.P.1			√								√			
Brick&Fence.B.1					√					√				
Brick&Fence.B.2	√				√					√				
Brick&Fence.B.3			√				√			√				
Brick&Fence.P.1					√					√				
Brick&Fence.P.2			√							√				

Table 7: Representation of orientation and structure for the subtextures of images in figure 2. The degrees of structure are: Perfectly Ordered (P.O), Almost Ordered (A.O.), Certain Order (C.O), Almost Random (A.R.) and Completely Random (C.R.).



Original image Dark blobs ($\sigma = 0.75$) Bright blobs($\sigma = 0.75$) Background($\sigma = 0.75$)

Dark blobs ($\sigma = 3$) Bright blobs($\sigma = 3$) Background($\sigma = 3$)

Figure 8: Blob segmentation and emerging patterns by filtering at different scales.

edges and the second one is that the sign of the pixels in all the regions of I' allow to state if the blob is bright, dark or if the corresponding area belongs to the image background.

The effects of the scale parameter on this method are shown in figure 8, where subtextures formed by small blobs are segmented by filtering with small scales and subtextures formed by emergent patterns are segmented by filtering with large scales.

The second step is to obtain global measurements on properies of the subtextures decomposed in the previous step. The method must detect not only existing blob structures but also emerging patterns from perceptual grouping of blobs. The Fourier transform is a global method, which perfectly suits our interests [15], because it emphasize the information of the underlying structure of the texture. In order to obtain information about the form and the relative position of each image subtexture we do the Fourier transform of the texture image multiplied by the masks we have obtained in the first step of the method, corresponding to blobs and background segmentations.

We have already established that there are three features we want to obtain from every subtexture: degree of structure, predominant orientations and scale. We propose to get all these measurement from the Fourier Spectrum of the texture.

Assuming that the Fourier transform is a decomposition of a function as an infinite sum of exponential functions with coefficients, and considering these coefficients are complex, we will only use their amplitude. Thus, a peak in the Fourier spectrum of a texture means the corresponding frequency is important in the image. In order to have an overview of all the important frequencies in the image, we will extract the significant peaks of the Fourier Spectrum and study their amplitude, position and stability. The form of these peaks will depend on how regular the structure is. This means we need to know not only the position and magnitude of the existing peaks, but also how sharp they are. In order to obtain all this information about the peaks in the Fourier spectrum, we will use a thresholding approach. For a fixed set of thresholds we calculate the area and the number of maxima we obtain. At the end of this process, we have all the information about the magnitude and sharpness of the peaks in the Fourier Spectrum.

At this point, we have got all the information necessary to begin the interpretation of subtexture properties. It will require knowing how these peaks are organized in the Fourier spectrum. The Hough transform detects the presence of parametrically representable groups of points in an image and it has been used to extract the underlying structure in the spectrum, if it exists. We will use it to find if peaks are grouped as straight lines or centered circles [23], from which basic measurements on subtexture properties will be derived.

From the information inferred from all the previous processes, let us to give an overview on how to describe the essential properties of a subtexture, that will be summed up in a six-dimensional representation, $(s_1, s_2, s_3, s_4, s_5, s_6)$:

Structure (s_1), indicates how structured the subtexture is, and is computed from the number of sharp peaks and the existence of straight lines in the spectrum.

Scale (s_2), represents the scale of the subtexture. Its computation will depend on whether the subtexture is structured or not. If it is, scale will be given by the most important frequency, and if it is not it will be derived from the radius of the most representative centered circle.

Orientation (s_3, s_4, s_5, s_6), represents the orientations of the four predominant directions. In case there are no predominant directions, all values will be 0, otherwise they will have the value of the corresponding orientation. They are deduced from the directions of the straight lines in the spectrum and the position of the peaks.

At this point we can completely characterize a texture, I, by a two-dimensional matrix, denoted as $\mathbf{R}$ and given by

$$\mathbf{R}_{\mathbf{N}\times\mathbf{6}} = (s_{ij}) \tag{3}$$

where N is the number of subtextures needed to represent texture I, where ith row corresponds to the feature vector of each the ith subtexture.

7 Summary

The main objective of this work has been to give a general blob-based texture representation. It has to allow to collect data on how textures are perceived by human subjects, without constraints due to usual closed-lab conditions. Most common experiments have been usually based on two different type of experiments, some based on preattentive segregation and the rest based on similarity judgements. In both approaches, the selection of a given set of image could condition their conclusions, since in any case we can not assure that a finite set of textures is enough representative of any texture. Moreover, it has been shown in other works that experiments based on attentive cognition processes can be affected by content-dependency effects.

The characteristics of this representation make us to suggest a web-based experiment to collect a large amount of experiments on different subjects and on wide sets of images since it can remain open for a long time and available from anywhere.

A large dataset on how textures are perceived would be an excellent tool to test algorithms for the computer vision community. It could improve current performance evaluation procedure based on building large image databases from dividing texture images in small parts.

Other interesting aspects could be the use of this data from a psychophysical point of view to infer perceptual representation spaces and their properties such as dimensionality or identification of relevant dimensions.

As a final remark we want to note that all this work has to be understood as a first proposal and some aspects still remain to be improved. Parts of the representation such as as those concerning the especification of emergent patterns needs to be better justified. This first approach has been mainly based on the results of previous experiments with 15 subjects. Although it is still incomplete we think it can be a useful tool towards a validation of computational representation of textures.

8 Acknowledgements

This work has been partially supported by the projects FEDER 2FD97-1800 and CICYT TIC2000-0382.

References

[1] J. Beck, K. Prazdny, and A. Rosenfeld. *Human and Machine Vision*, pages 1–38. Academic, New York, 1983.

[2] J. Beck, A. Sutter, and R. Ivry. Spatial frequency channels and perceptual grouping in texture segregation. *Computer Vision, Graphics, and Image Processing*, 37:299–325, 1987.

[3] J. Bergen. Theories of visual texture perception. In D. Regan, editor, Vision and Visual Dysfunction, volume 10B, pages 114–134. Macmillian, New York, 1991.

[4] J.R. Bergen and E.H. Adelson. Early vision and texture perception. *Nature*, 333:363–364, 1988.

[5] J.R. Bergen and M.S. Landy. *Computational Models of Visual Processing*, chapter Computational Modeling of Visual Texture Segregation, pages 253–271. The MIT Press, 1991.

[6] R. Boynton and C.X. Olson. Locating basic colors in the OSA space. *Color Research and application*, 12(2), 1987.

[7] P. Brodatz. *Textures.* New York: Dover, 1966.

[8] P.S.P. Wang C.H. Chen, L.F.Pau, editor. *The Handbook of Pattern Recognition and Computer Vision*, chapter Texture Analysis, pages 207–248. World Scientific Publishig Co., 1998.

[9] K. Dana, S. Nayar, B. Ginneken, and Koenderink J. Reflectance and texture of real-world surfaces. In *Proceedings of the CPR97*, pages 151–157, 1997.

[10] S. Edelman. *Representation and Recognition in Vision.* M.I.T. Press, 1999.

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

[12] G.L. Gimel'farb and A.K. Jain. On retrieving textured images from an image database. *Pattern Recognititon*, 29:1461–1483, 1996.

[13] L.van Gool, P. Dewaele, and A. Oosterlinck. Survey: Texture analysis anno 1983. *Computer Vision, Graphics and Image Processing*, 29:336–357, 1985.

[14] R. Gurnsey and D. Fleet. Texture space. *Vision Research*, 41:745–757, 2000.

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

[16] L. Harvey and M.Gervais. Internal representation of visual texture as the basis for the judgment of similarity. *Journal of Experimental Psychology: Human Perception and Performance*, 7(4):741–753, 1981.

[17] C. Heaps and S. Handel. Similarity and features of natural textures. *Journal of Esperimental Psychology: Human Perception and Performance*, 25(2):299–320, 1999.

[18] A. Jain and F. Farrokhnia. Unsupervised texture segmentation using gabor filters. *Pattern Recognition*, 24(12):1167–1186, 1991.

[19] B. Julesz. Experiments in the visual perception of texture. *Scientific American*, pages 34–44, 1978.

[20] B. Julesz. Textons, the elements of texture perception, and their interactions. *Nature*, 290:91–97, 1981.

[21] B. Julesz. Texton gradients: The texton theory revisited. *Biological Cybernetics*, 54:245–251, 1986.

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

[23] V.F. Leavers. *Shape Detection in Computer Vision Using the Hough Transform*. Springer-Verlag, 1992.

[24] T. Lindeberg. *Scale-Space Theory in Computer Vision*. Kluwer Academic Publishers, 1994.

[25] J. Malik and P. Perona. Preattentive texture discrimination with early vision mechanisms. *Journal of the Optical Society of America*, 7:923–932, 1990.

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

[27] D. Marr. *Vision*. W. H. Freeman and Company, 1982.

[28] P. Ohanian and R. Dubes. Performance evaluation for four classes of textural features. *Pattern Recognition*, 25(8):819–833, 1992.

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

[30] A.R. Rao and Gerald L. Lohse. Identifying high level features of texture perception. *CVGIP: Graphical Models and Image Processing*, 55:218–233, 1993.

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

[32] D. Sagi. *Channels in the visual nervous system: Neurophysiology, Psychophysics and Models*, chapter Spatial filters in texture segmentation tasks, pages 397–424. London: Freund Publishing House, 1991.

[33] J. Sturges and T.W.A Whitfield. Locating basic colours in the Munsell space. *Color Research and application*, 20(6), 1995.

[34] H. Tamura, S. Mori, and T. Yamawaki. Textural features corresponding to visual perception. *IEEE Trans. on System, Man and Cybernetics*, SMC-8:460–473, 1978.

[35] Vision Texture. http://www-white.media.mit.edu/vismod/imagery/VisionTexture/vistex.html.

[36] M. Vanrell, X. Roca, and J. Vitrià. A multidimensional scaling approach to explore the behavior of a texture perception algorithm. *Machine Vision and Applications*, pages 262–271, 1997.

[37] H. Voorhees. Finding texture boundaries in images. Technical Report 968, MIT Artificial Intelligence Laboratory, 1987.

[38] B. Wandell. The synthesis and analysis of color images. *IEEE PAMI*, 9(1), 1987.

[39] H. Wechsler. Texture analysis - a survey. *Signal Processing*, 2:271–282, 1980.

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