# A four-dimensional texture representation space

Maria Vanrell and Jordi Vitrià
Computer Vision Center - UAB - 08193 Bellaterra (Barcelona)

Abstract

In this work we present a first result towards the definition of a general computational texture representation. We propose some properties that a general representation should verify. Taking these properties as the basis, we define a representation based on a computational model of the human vision system. An exploratory analysis followed by a confirmatory step take us to formulate a four-dimensional texture representation space where any texture can be represented considering the attributes of its own textons.

Key Words : Texture representation, texton attributes, low-dimensional representation.

# A four-dimensional texture representation space

Maria Vanrell and Jordi Vitrià
Computer Vision Center - UAB - 08193 Bellaterra (Barcelona)

Abstract

In this work we present a first result towards the definition of a general computational texture representation. We propose some properties that a general representation should verify. Taking these properties as the basis, we define a representation based on a computational model of the human vision system. An exploratory analysis followed by a confirmatory step take us to formulate a four-dimensional texture representation space where any texture can be represented considering the attributes of its own textons.

Key Words: Texture representation, texton attributes, low-dimensional representation.

## 1 Introduction

Texture is an important cue in visual tasks as segmentation, classification and shape extraction. There are many works which have focused in the texture perception problem and there are different approaches to deal with it [2, 14, 9]. The objective can be seen as to find a computational description with the ability to represent any texture and to derive from it the mechanisms that allows to solve specific applications.

This objective has been reached for other important visual stimulus as color. There exists different representations with interesting properties. These representations has standardized the way in which the problems with color in computer vision are treated. A similar objective has been undertaken with the problem of representing the shape of 3D objects [1].

Considering this point of view there is an interesting result about the texture perception problem: R. Rao et al [6, 7] have deduced a three-dimensional space of characteristics with a metric which agrees with the texture perceptual space of the human visual system (HVS). The axes of this space have been associated to higher order perceived features as granularity, contrast, repetitiveness and direccionality.

The above result encourages us about the possibility to define a computational representation for any texture that could accomplish interesting properties as generality and low-dimensionality. Following this idea, we adopt as the objective of this work to find a computational representation of textures. The work has been organized in the following steps. Firstly we define the properties that a texture representation should fulfil. Secondly, we define a texture representation based on a computational model for texture perception. Finally, we demonstrate that this complex representation defines a four-dimensional texture representation space with a perceptual interpretation for each axis.

## 2 Properties of a texture representation space

Different approaches in computer vision have given different computational texture descriptions:

Structural approach Texture is described by the set of shapes of the basic elements and the placement rules that organize these shapes in the texture.

Mathematical modelization Texture is considered as the realization of a concrete mathematical model, hence it is defined by the model parameters.

Measure-based approach Texture is represented by a vector of characteristics measuring certain higher order perceived features, such as: granularity, repetitiveness, direccionality, etc.

Filtering approach Texture is described by the set of channels corresponding to the responses of convolving a set of filters with the image. Each channel gives the tunned response of the image to a certain type of filter.

The main constraint of the two first approaches is generality. Any texture can not be expressed by the localization of its shapes, and there not exists a general model which provides a description of any texture. The measure-based approach has been widely used in several applications. The main problem of this approach is given by the fact that measures present certain degree of ambiguity and a non unique mechanism to compute them. The last approach presents an important complexity given by the high number of responses that have to be incorporated to be a general mechanism.

From the analysis of the previous approaches now we are going to give the properties that a general texture representation should fulfil. Generality, it allows to represent every texture. Non-ambiguity, it should be based on photometric or geometric characteristics of the texture regions and not on higher order measures. Computational efficiency, it should be computed by an efficient algorithm. Low-dimensionality, the representation space should be based on a low dimensional space. Perceptual interpretation, it should be interesting to obtain a perceptual interpretation of each axis to be able to derive practical applications. With a similarity metric like the HVS, it should be interesting to establish a second order isomorphism between the representation space and the perceptual space defined by the HVS.

Now we are going to define a representation based on the last approach and we will show its feasibility as a general texture representation fulfilling the suggested properties.

## 3 A multichannel representation

We have defined a texture representation based on a computational model of human preattentive texture perception. The model selected has been defined by J. Malik and P. Perona in [5] as a general model for texture segmentation. We define a texture representation taking the responses of the inhibition network denoted as $PIR_i$ for the $i$ channel [12]. For an image with an unique texture the responses after the inhibition network present an homogeneous gray-level which depends on how the filter has been tunned to the local features of the texture.

Definition 1 For a $N \times N$ image with a texture $t$ we define its representation, $r$, as

$$r(t) = (v_1, ..., v_n) \quad where \quad v_i = \sum_{(x,y)\in\{0,\ldots,N\}^2} \frac{PIR_i(x,y)}{N^2} \tag{1}$$

and $n$ depends on the number of channels in the model. The parameters recommended by the authors imply $n = 192$.

This representation has demonstrated a good behavior in representing any texture. It has been applied to represent images in a large database giving satisfactory results on retrieval operations over these large sets of images [12].

Provided that, the representation is based on a model of preattentive perception we can assure that it is considering the geometrical and photometric attributes of the texton elements. In this sense, there not exists any ambiguity about what is it representing.

A fast algorithm to compute this representation has been defined in [10, 12], by introducing morphological operations in the inhibition steps, in order to reduce the computational complexity.

| Dimension | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| Set 1 | 0.37402 | 0.19609 | 0.10843 | 0.05063 | 0.02685 |
| Set 2 | 0.35341 | 0.15683 | 0.08905 | 0.04881 | 0.03614 |

Table 1: Stress values obtained by MDS over two sets of images to a determined dimension.

Heretofore, we can state that the defined representation accomplishes the first three properties. Now we are going to see how the representation behaves with respect to the others.

# 4 Axes of the representation space

From definition 1 we obtain a representation space of 192 dimensions. The question that we want to respond is can be the structure of this space maintained in a low-dimensional space?; if it is true then which is the underlying dimension of this representation space? and which are the axes of this low-dimensional space?.

In order to respond to these questions we have used the Multidimensional Scaling (MDS) method. It has been developed to analyze and visualize data in such a way that, configurations of points in a space of dimension $m$ can be seen in a space of dimension $k$, where $k < m$, preserving the interpoint distances in a monotonic sense [8, 4]. To measure which error is made in the scaling process it has been defined the stress measure, denoted as $S$.

The first two questions imply to find the true dimensionality of the representation space. It can be done from the analysis of the stress values obtained from scaling points of the representation space to some determined dimensions. The inability to do this for the set of all possible textures take us to select a subset of textures with a great variety of characteristics and to consider that it will represent the set of all textures. This assumption has also been made by other authors [7].

We have obtained that the true dimension of the representation space is four. In table 1 we can see that the stress values are importantly reduced when the space is scaled to a four dimensional space. The same result has been given from two different sets of textures.

Using the MDS method on the same images we are going to answer the last question in two steps:

1. Exploring the representation space using parametric textures and deriving a first hypothesis about variations on each space.

2. Confirming the previous hypothesis by a linear multiple regression analysis over a set of natural images that we will assume as a general set of textures.

## 4.1 Hypothesis 1

The representation used is mainly based on the Julesz's texton theory [3]. Therefore, it is considering densities of texton attributes. For this reason we have considered an image model based on the bars and blobs of the image and their attributes. This image model has allowed to define a parametric texture space from which we can synthesize images controlling the attributes of their bars and blobs.

A texture in the parametric space consists of a repetition of a single texton (bar or blob) with specific parameters. Therefore, an image is given by the texton parameters, which are the contrast, the length, the size and the orientation. For any image in the parametric space we do not regard on texton localization as the representation does not do.

To explore the representation space we have studied how variations of certain attributes on images of the parametric space affect their representations. Considering all possible variations we have defined a first hypothesis about the behavior of the representation. The details of this work can be seen in [11, 12]. The results can be summarized in the following hypothesis.

Hypothesis 1 The representation, $r$, of a single texture, $t$, that is, with a unique type of texton, is given by its attributes $s$, $l$, $c$ and $\theta$, which defines its size, length, contrast and orientation, respectively. The representation of this texture in a four dimensional space is given by the following four coordinates:

$$b(t) = (b_1(s,l,|c|)\cos(2\theta), b_2(s,l,|c|)\sin(2\theta), sgn(c)\cdot b_3(|c|), b_4(|c|)) \; b_i : monotonic\ functions \tag{2}$$

## 4.2 Hypothesis 2

In order to extend the hypothesis 1 for the general case we should demonstrate that the representation of any texture accomplishes this hypothesis.

Toward this objective, we need to describe any texture by the attributes of its textons. We could apply a segmentation approach followed by the characterization of the attributes as it has been done in [13]. However, we have preferred to construct a parametric interpretation of the computed representation.

The interpretation considers the type of filter defining a channel, that is, the orientation and sign associated to a given channel. On the other hand, for a set of channels corresponding to one filter the interpretation will consider the scale and the contrast of the better tunned channel.

Hence, we are able to give an interpretation of a texture, $t$, in terms of the texton attributes that it contains.

$$t \rightarrow \{t_1, ..., t_{16}\} \rightarrow \{(s_1, c_1, \theta_1), ..., (s_{16}, c_{16}, \theta_{16})\} \tag{3}$$

The number of simple textures is due to the fact that the model uses eight filters, two isotropic filters and a non-isotropic one with six different orientations. For this type of filters the model separates positive and negative part of the responses obtaining sixteen different groups of channels. To compute the different parameters we will redefine the representation vector as

$$r(t) = (v_1, \cdots, v_{192}) \;\longrightarrow\; r'(t) = (v_1^1, \cdots v_{12}^1, \cdots, v_1^{16}, \cdots, v_{12}^{16}) \tag{4}$$

where $v^k$ represents positive respones for $1 \leq k \leq 8$ and negative responses for $8 < k \leq 16$.

For each group of channels we compute the values of $s_k$, $c_k$ and $\theta_k$, where $k \in \{1, ..., 16\}$. Every group defines the parameters of a simple texture. They are computed by the following expressions:

$$c_k = \delta(k) \max_{i=1\cdots 12} v_i^k \quad : \quad \delta(k) = \begin{array}{ll} 1 & 1 \leq k \leq 8 \\ -1 & 9 \leq k \leq 16 \end{array} \tag{5}$$

$$s_k = s \; : \; v_s^k = \max_{i=1\cdots 12} v_i^k \tag{6}$$

$$\theta_k = \begin{array}{ll} \emptyset & \text{If } k \in \{1,2,9,10\} \\ 0^\circ, 30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ & \text{If } k \in \{3,11\},\{4,12\},\{5,13\},\{6,14\}, \\ & \{7,15\},\{8,16\} \text{ respectively} \end{array} \tag{7}$$

(8)

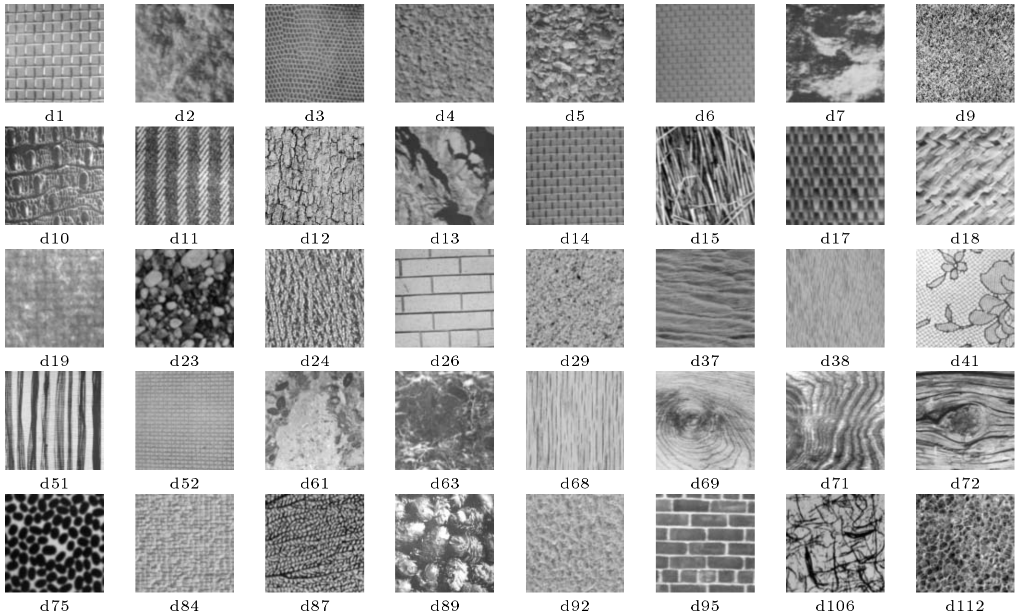


Figure 1: Selected images to confirm the hypothesis

Given this parametric interpretation of any texture we will suppose the sum of representations as the way to introduce any texture in the representation space. This fact is included in the following hypothesis:

Hypothesis 2 The representation of a texture $t$ can be seen as the sum of the representations of a set of simple textures $\{t_1, ..., t_{16}\}$. Every simple texture, $t_i$, is given by its parameters $(c_i, s_i, \theta_i)$ and its representation in a four-dimensional space is given by the hypothesis 1, which implies that

$$b(t) = \left(\sum_{i=1}^{16} b_1(t_i), \sum_{i=1}^{16} b_2(t_i), \sum_{i=1}^{16} b_3(t_i), \sum_{i=1}^{16} b_4(t_i)\right) \tag{9}$$

## 4.3 Confirmation of the hypothesis

As a confirmatory step of the two previous hypothesis we will carry out a multiple linear regression analysis. That is, we will show that the variables $b_1$, $b_2$, $b_3$ and $b_4$, introduced by the hypothesis 1 and computed for any texture considering the hypothesis 2, correspond with important axes in the representation space.

In order to do this, we will first compute the representation of the set of textures given in figure 1. We will assume that this set of 40 images represent a complete sampling of the space of textures, a similar assumption has been justified in [7].

A linear regression analysis over the set of representations allows to find the straight line in the space that better correlates with the values of a certain variable. In our case variables will be given by the values of each coordinate, $b_i$, of the four-dimensional space. Any texture, $t_j$, of the defined set will be given in the representation space as $r(t_j) = x_j$.

The regression analysis has to find the lines $\bar{b}_i$ that better approximate each variable:

$$b_i \cong \bar{b}_i = B_0 + B_1 x_1 + \cdots + B_{40} x_{40} \tag{10}$$

The results obtained of computing this line for each variable present the following Pearson correlation coeficients, $R = 0.875, 0.810, 0.660, 0.938$, for $b_1$, $b_2$, $b_3$ and $b_4$, respectivelly.

From these results we can state that variables $b_1$, $b_2$ and $b_4$ can be considered as important axes of the representation space. On the other side, variable $b_3$ representing the combination of the sign of the blobs should be ameliorated to be an important axis in the representation space.

The importance of this study is that we have showed that from a general representation we can define a low-dimensional space that can represent any texture. Moreover, we have identified the perceptual meaning of each axis in this space, which will allow to derive certain application with the ability to control certain texture parameters using this representation.

# 5 Conclusions

The objective of this work has been to give a first approximation to the problem of finding a computational representation for the texture visual cue. The same objective has been treated by different authors with other cues. We have introduced some properties as those that any interesting representation should verify.

The computational representation that we propose is based on a computational model of the human vision system. Then, we have demonstrated that it verifies the required properties. Its feasibility as a general texture representation has been showed in some experiments of retrieving images in a large database of textures [12].

The defined representation space can be understood as a four-dimensional space with a perceptual interpretation for each axis. The interpretation is given in terms of the textons attributes of the image. We have proved the validity of this interpretation with a linear regression analysis over a set of images that we have assumed as a general sample.

A better knowledge of these dimensions will take us to know the exact structure of the representation space. Then, we will be able to show if there exists a second order isomorphism between this space and the perceptual space of the HVS. This fact would validate the behavior of the representation with respect to HVS.

# References


[1] F. Cutzu and S. Edelman. Exploration of shape space. Tech. report, Weizmann Institute, 1995.

[2] L.van Gool et al. Survey: Texture analysis anno 1983. CVGIP,29:336–357,1985.

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

[4] J.B. Kruskal and M. Wish. Multidimensional Scaling. Sage University Paper, 1978.

[5] J. Malik and P. Perona. Preattentive texture discrimination with early vision mechanisms. JOSA, 7:923–932,1990.

[6] A.R. Rao and Gerald L. Lohse. Identifying high level features of texture perception. CVGIP:GMIP, 55:218–233, 1993.

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

[8] R.N. Shepard. The analysis of proximities: multidimensional scaling with an unknown distance function. i and ii. Psychometrika, 27:125–140,219–246, 1962.

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

[10] M. Vanrell and J. Vitrià. 3 3 decomposition of circular structuring elements. In ICIP, volume 3, pages 5–8. IEEE, 1996.

[11] M. Vanrell, J. Vitrià, and X. Roca. A multidimensional scaling approach to explore the behavior of a texture perception algorithm. Machine Vision and Applications, 1996. To be published.

[12] M. Vanrell. Identificació de les dimensions d'un espai de representació de textures basat en un model de percepció preatentiva. PhD thesis, Universitat Autònoma de Barcelona, 1996.

[13] H. Voorhees and T. Poggio. Computing texture boundaries from images. Nature, 333:364–367,88.

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