

Manuscript Number: IMAVIS-D-16-00270R2

Title: Extended three-dimensional rotation invariant local binary  
patterns

Article Type: Full Length Article

Keywords: Local binary patterns (LBP); Three-dimensions; Rotation  
invariance; Texture classification

Corresponding Author: Mr. Leonardo Citraro, MSc.

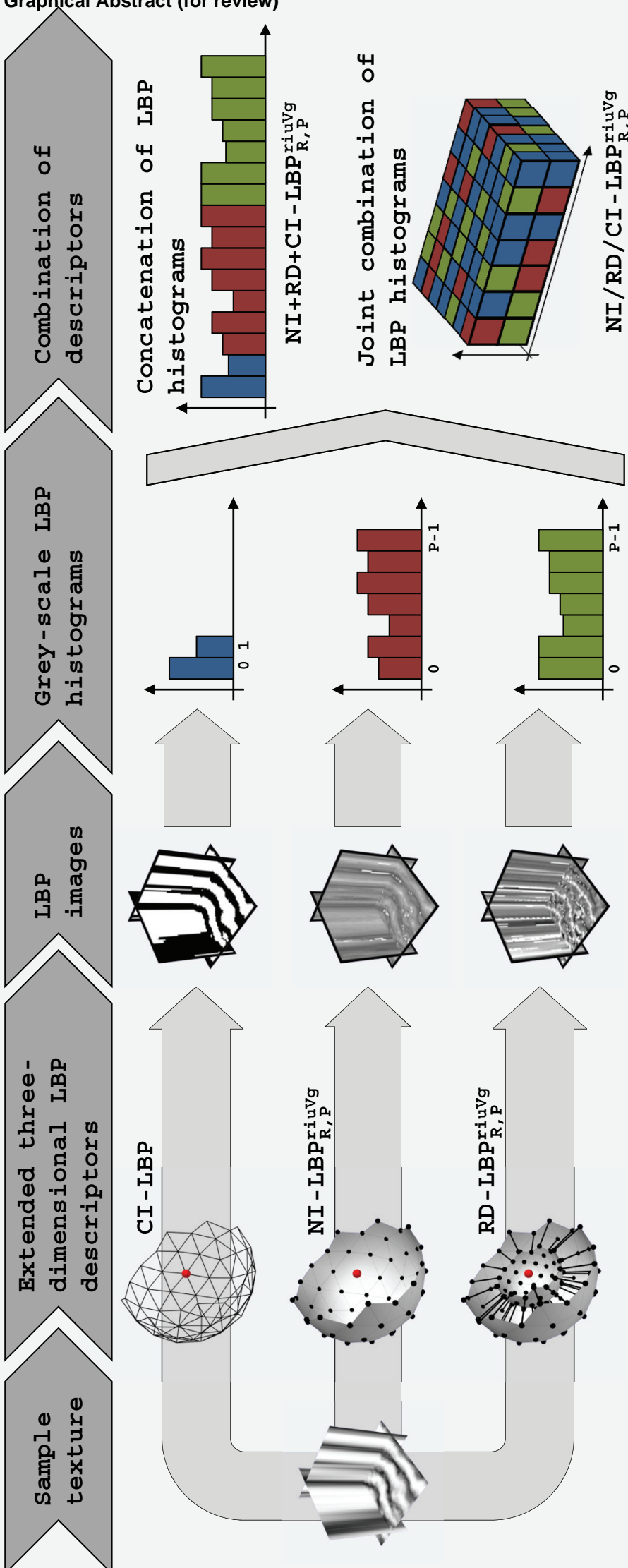
Corresponding Author's Institution: University of Southampton

First Author: Leonardo Citraro, MSc.

Order of Authors: Leonardo Citraro, MSc.; Sasan Mahmoodi, Professor, Phd;  
Angela Darekar, Phd; Brigitte Vollmer, Professor, Phd

Abstract: This paper presents a new set of three-dimensional rotation invariant texture descriptors based on the well-known local binary patterns (LBP). In the approach proposed here, we extend an existing three-dimensional LBP based on the region growing algorithm using existing features developed exquisitely for two-dimensional LBPs (pixel intensities and differences). We have conducted experiments on a synthetic dataset of three-dimensional randomly rotated texture images in order to evaluate the discriminatory power and the rotation invariant properties of our descriptors as well as those of other two-dimensional and three-dimensional texture descriptors. Our results demonstrate the effectiveness of the extended LBPs and improvements against other state-of-the-art hand-crafted three-dimensional texture descriptors on this dataset. Furthermore, we prove that the extended LBPs can be used in medical datasets to discriminate between MR images of oxygenated and non-oxygenated brain tissues of newborn babies.

This paper presents a new set of three-dimensional rotation invariant texture descriptors based on the well-known local binary patterns (LBP). In the approach proposed here, we extend an existing three-dimensional LBP based on the region growing algorithm using existing features developed exquisitely for two-dimensional LBPs (pixel intensities and differences). We have conducted experiments on a synthetic dataset of three-dimensional randomly rotated texture images in order to evaluate the discriminatory power and the rotation invariant properties of our descriptors as well as those of other two-dimensional and three-dimensional texture descriptors. Our results demonstrate the effectiveness of the extended LBPs and improvements against other state-of-the-art hand-crafted three-dimensional texture descriptors on this dataset. Furthermore, we prove that the extended LBPs can be used in medical datasets to discriminate between MR images of oxygenated and non-oxygenated brain tissues of newborn babies.



Extended three-dimensional rotation invariant Local Binary Patterns  
Image and Vision Computing IMAVIS-D-16-00270  
Highlights (bullet points)

Leonardo CITRARO, ldo.citraro@gmail.com

November 27, 2016

## 1 Highlights

---

- A new set of three-dimensional fully rotation invariant LBP descriptors is proposed.
- Proven utility of the third dimension in local binary patterns.
- Improvements against other state-of-the-art 3D texture descriptors.
- Application to a clinical dataset of susceptibility-weighted MR brain images.

# Extended Three-Dimensional Rotation Invariant Local Binary Patterns

Leonardo Citraro<sup>a,\*</sup>, Sasan Mahmoodi<sup>a</sup>, Angela Darekar<sup>b</sup>, Brigitte Vollmer<sup>c</sup>

<sup>a</sup>*School of Electronics and Computer Science, Building 1, Southampton University, Southampton SO17 1BJ, UK*

<sup>b</sup>*Department of Medical Physics, University Hospital Southampton NHS Foundation Trust, Southampton, SO16 6YD, UK*

<sup>c</sup>*Clinical Neurosciences, Clinical and Experimental Sciences, Faculty of Medicine, University of Southampton; University Hospital Southampton NHS Foundation Trust, Southampton, SO16 6YD, UK*

---

## Abstract

This paper presents a new set of three-dimensional rotation invariant texture descriptors based on the well-known local binary patterns (LBP). In the approach proposed here, we extend an existing three-dimensional LBP based on the region growing algorithm using existing features developed exquisitely for two-dimensional LBPs (pixel intensities and differences). We have conducted experiments on a synthetic dataset of three-dimensional randomly rotated texture images in order to evaluate the discriminatory power and the rotation invariant properties of our descriptors as well as those of other two-dimensional and three-dimensional texture descriptors. Our results demonstrate the effectiveness of the extended LBPs and improvements against other state-of-the-art hand-crafted three-dimensional texture descriptors on this dataset. Furthermore, we prove that the extended LBPs can be used in medical datasets to discriminate between MR images of oxygenated and non-oxygenated brain tissues of newborn babies.

**Keywords:** Local binary patterns (LBP), Three-dimensions, Rotation invariance, Texture classification

---

## 1. Introduction

### 1.1. Context

Texture analysis is a key topic in image processing and computer vision, playing an important role in medical and industrial applications. Texture is a fundamental property of physical objects that, when captured on image, may suffer from noise, illumination changes, occlusions, viewpoints or scale variations. Producing high discriminant descriptors invariant with respect to grey-scale changes as well as to rotation is essential in order to solve many real-world problems. With the advent of three-dimensional images, textures have become more complex and full of information leading to extremely interesting applications. Although two-dimensional texture analysis methods have become very powerful, in three-dimension, these approaches may result in the loss of important information with consequent non-optimal results. An appropriate characterisation of textures in three-dimensions is therefore crucial for the development of state-of-the-art systems as for cases such as the analysis of MR images (Kovalev et al., 2001) or modelling and recognition of 3D biomedical textures (Depeursinge et al., 2014; Majtner and Svoboda, 2014). Among the most popular two-dimensional texture methods, the local binary patterns (LBP) have gained a great deal of attention in many applications and research studies in the last decade (Ahonen et al., 2006; Liu et al., 2012; Ojala et al., 2002; Sorensen et al., 2010). In contrast,

the three-dimensional implementations of local binary patterns are very few and often do not fully match the necessary requirements. Zhao and Pietikäinen (2006, 2007b,a) introduce for the first time the concept of three-dimensional local binary pattern by proposing two novel approaches to analyse dynamic textures. The first method (*VLBP*) consists of stacking multiple classic LBP operators on top of each other forming a spiral whereas the second method (*LBP – TOP*) consists in placing three descriptors in an orthogonal fashion. Paulhac et al. (2008) propose a three-dimensional fully rotation invariant LBP descriptor based on a region growing algorithm and the so called uniform patterns. Fehr and Burkhardt (2008) attempted to classify 3D volume data exploiting the LBP method and the spherical harmonics (SH). In a similar way, Banerjee et al. (2013) propose a descriptor by exploiting the spherical harmonics but unlike Fehr and Burkhardt (2008) the rotation invariance is obtained without searching for the minimum correlation over all angles. Another interesting method is proposed by Liu et al. (2011) who developed a novel three-dimensional fully rotation invariant LBP based on unique rotation, reflection and translation invariant patterns. This method is based on the extraction in advance of all possible congruent patterns which are then clustered in order to identify all the unique shapes defining a texture.

In addition to these papers, there is a considerable number of publications related to local binary patterns and the work we present here. Majtner and Svoboda (2014) present a comparison of 2D and 3D texture descriptors on biological data. The volume LBP *VLBP* and the *LBP – TOP*, as well as four variants of non-rotation invariant 3D LBP built from four Platonic solids frames are investigated in their paper. The

---

\*Corresponding author

Email addresses: ldo.citraro@gmail.com (Leonardo Citraro), sm3y07@soton.ac.uk (Sasan Mahmoodi), angela.darekar@uhs.nhs.uk (Angela Darekar), b.vollmer@soton.ac.uk (Brigitte Vollmer)

results obtained in Majtner and Svoboda (2014) demonstrate good classification accuracy for all methods investigated in their paper on a dataset of fluorescence microscopy images. Surface descriptors are also becoming beneficial in many applications such as face-based recognition and multi-view stereo systems as well as 3D scanning and photometric data (Bayramoglu et al., 2013; Werghi et al., 2013, 2015a,b; Biasotti et al., 2016). Bayramoglu et al. (2013) propose oriented based local binary pattern descriptors to address the task of detecting facial action units in 3D face data. In a similar way, Werghi et al. (2013, 2015a,b) present a novel approach for computing local binary pattern on triangular mesh manifolds addressing a face recognition problem on three-dimension scans instead of 3D depth images (Li et al., 2005; Huang et al., 2006; Huynh et al., 2012). Also, Biasotti et al. (2016) present a remarkable comparative study of several different state-of-the-art methods for the classification of three-dimensional texture models.

In this paper our contribution is to propose a new set of three-dimensional rotation invariant descriptors based on the local binary patterns. We make use of the concepts of intensity-based (*CI* and *NI*) and difference-based (*RD*) features proposed by Liu et al. (2012) in order to improve the existing three-dimensional LBP descriptors proposed by Paulhac et al. (2008). We conduct extensive experiments on the proposed descriptor in order to compare their performance with existing state-of-the-art two-dimensional and three-dimensional methods. The dataset employed here is a set of three-dimensional synthetic images constructed by interpolation (Paulhac et al., 2009).

### 1.2. Local binary patterns

The local binary pattern (LBP), first introduced by Harwood et al. (1995) and Ojala et al. (1996), is a method that encodes the local texture information in a binary fashion by simply sampling and weighting the neighbouring pixels. In their later work, Ojala et al. (2002) propose a circular and rotation invariant LBP composed of  $P$  equally spaced grey-scale pixels ( $g_{(R,P)}$ ) arranged on a circle of radius  $R$ . These values  $g_{(R,P)}$  try to replicate the local texture at some point of the image, therefore, their values depend upon the neighbouring pixels. In details, each pixel value  $g_{(R,P)}$  is assigned a grey-scale value based on the bilinear interpolation of the four closest image pixels. Successively, the LBP pixels are sampled (threshold-ing process) and concatenated starting (arbitrarily) from the least significant  $g_{(R,P=0)}$  to the most significant  $g_{(R,P-1)}$  in order to form a binary number. The resulting value, also called LBP code, characterizes the local texture. Formally, the circular non-rotation invariant LBP is defined as:

$$s(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases} \quad (1)$$

$$LBP_{R,P} = \sum_{p=0}^{P-1} s(g_{(R,p)} - g_{(0,0)}) 2^p \quad (2)$$

where  $s(x)$  is the sampling (threshold) function over  $x$ . Ojala et al. (2002) also propose a method to achieve the rotation invariance which allows to obtain consistent results even with rotated images. This property is obtained by applying a simple

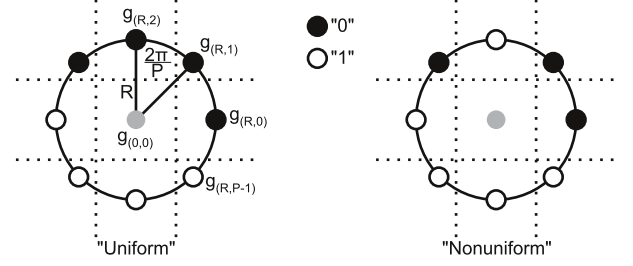


Figure 1: Example of "uniform" and "nonuniform" geometrical representation of the circular LBP presented in (Ojala et al., 2002). A "uniform" LBP is defined by two or fewer bitwise 0/1 transitions. The pattern on the left produces an LBP code of four whereas the pattern on the right is assigned the value nine.

circular bit-wise right shifts until the minimum possible binary value is reached:

$$LBP_{R,P}^{ri} = \min\{ROR(LBP_{R,P}, i) | i = 0, 1, \dots, P-1\} \quad (3)$$

where  $ROR(x, i)$  is the circular bit-wise right shifts operator and the superscript  $ri$  indicates the rotation invariance capability. In addition to the rotation invariant LBP, Ojala et al. (2002) have discovered that many LBP codes represent fundamental properties of textures as lines, end of lines, flat regions or edges. These particular codes share the same specific pattern, the number of spatial transitions (bitwise 0/1 changes of the LBP pixels after sampling) on the circular frame. This pattern, also called uniformity measure, is defined as:

$$U(LBP_{R,P}) = |s(g_{(R,P-1)} - g_{(0,0)}) - s(g_{(R,0)} - g_{(0,0)})| + \sum_{p=1}^{P-1} |s(g_{(R,p)} - g_{(0,0)}) - s(g_{(R,p-1)} - g_{(0,0)})| \quad (4)$$

where  $U(LBP_{R,P}) \leq 2$  is the definition of a "uniform" LBP and  $U(LBP_{R,P}) > 2$  "non-uniform". Thus, Ojala et al. (2002) have proposed a new descriptor  $LBP_{R,P}^{riu2}$  that encodes the local texture information using only  $P+1$  values:

$$LBP_{R,P}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_{(R,p)} - g_{(0,0)}) & \text{if } U(LBP_{R,P}) \leq 2 \\ P+1 & \text{otherwise} \end{cases} \quad (5)$$

Figure 1 shows an example of "uniform" LBP (left) that produces a code of four whereas the "nonuniform" LBP (right) is assigned the value  $P+1$ .

### 1.3. Extended local binary patterns

Liu et al. (2012) present a new set of improved LBP descriptors exploiting two types of features (pixel intensities and differences). The first intensity-based LBP proposed by Liu et al. (2012) is the *CI-LBP* which encodes the local contrast information using the mean of the whole image:

$$CI-LBP = s(g_{(0,0)} - \mu_r) \quad (6)$$

where  $\mu_r$  is the grey scale mean of the whole image. The second intensity-based feature proposed in Liu et al. (2012) is the *NI-LBP*. It encodes the texture by sampling the neighbouring pixels at the mean of the interpolated pixels.

$$NI-LBP_{R,P}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(g_{(R,p)} - \mu) & \text{if } U(NI-LBP_{R,P}) \leq 2 \\ P+1 & \text{otherwise} \end{cases} \quad (7)$$

where  $\mu = \frac{1}{P} \sum_{p=0}^{P-1} g_{(R,p)}$ . This descriptor has many merits such as the increased encoding capacity and the robustness. It is able to preserve weak edges, to discriminate between different homogeneous regions and it is less affected by the noise in the images compared to the classic  $LBP_{R,P}^{riu2}$ . Finally, the difference-based descriptor proposed in Liu et al. (2012) is the  $RD-LBP_{R,P}^{riu2}$ . In this descriptor the radial difference between two circular patterns with different radius is computed.

$$RD-LBP_{R,P,\delta}^{riu2} = \begin{cases} \sum_{p=0}^{P-1} s(\Delta_{\delta}^{Rad}) & \text{if } U(RD-LBP_{R,P,\delta}) \leq 2 \\ P+1 & \text{otherwise} \end{cases} \quad (8)$$

where  $\Delta_{\delta}^{Rad} = g_{(R,p)} - g_{(R-\delta,p)}$  and  $\delta$  is an integer defining the radial displacement between the external circle and the internal circle. For instance,  $RD-LBP_{3,24,1}^{riu2}$  computes the radial difference between an LBP with  $R = 3$  and  $P = 24$  and another LBP with  $R = 2$  and  $P = 24$ .

#### 1.4. Three-dimensional local binary patterns

Paulhac et al. (2008) propose a three-dimensional rotation invariant local binary pattern exploiting the concept of "uniform" LBP without passing through a binary representation of the voxels composing the LBP frame. The choice of developing a rotation invariant descriptor was inevitable due to the intractable number of possible unique codes a spherical LBP can have. In two dimensions, a non-rotation invariant LBP with  $P = 8$  has  $2^8$  possible resulting codes. In three dimensions, this number increases dramatically. The descriptor is composed of a series of concentric circles at different heights forming a sphere. As with the classic circular LBP, this three-dimensional descriptor is defined as:

$$LBP_{R,P'}^{riuV} = \begin{cases} \sum_{p=0}^{P'-1} s(g_{(R,p)} - g_{(0,0,0)}) & \text{if } U(LBP_{R,P'}) \leq V \\ P'+1 & \text{otherwise} \end{cases} \quad (9)$$

where  $P' = (S-1) \cdot P + 2$  defines the number of voxels on the spherical frame,  $S$  defines the number of circles,  $P$  defines the number of equally spaced voxels per circle and  $V$  is the threshold that determines if the LBP is "uniform" or not. Their method cleverly searches for the uniform conglomerates on the surface of the sphere without passing from a binary representation of the texture. A region growing algorithm groups similar binary voxels to form regions. If the number of regions is less than or equal to  $V$ , the LBP is defined as "uniform" otherwise it is defined as "nonuniform". Furthermore, they have decided to increase the number of regions  $V$  from two to three due to the small number of "uniform" LBP produced. Nevertheless, due to the relaxed constraint ( $V = 3$ ), some local binary patterns can have the same code despite the different texture. Figure 2 shows two different textures having the same LBP code.

## 2. Our proposed three-dimensional descriptors

### 2.1. Uniform spherical structure

An optimal rotation invariant LBP should have the same identical geometrical structure under all points of view. Such

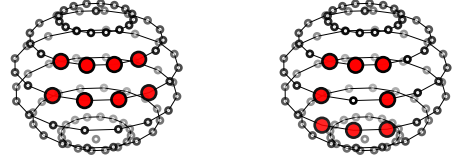


Figure 2: Example of two three-dimensional local binary patterns ( $S = 7, P = 16$ ) proposed by Paulhac et al. (2008) having the same LBP code of eight.

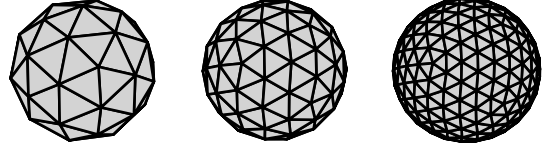


Figure 3: Geodesic spheres derived from the icosahedron. From left to right: 42, 92 and 252 vertices

a geometry on a sphere is hard to obtain, however, there are structures that accomplish our requirements. A geodesic grid (Kenner, 2003) is a technique that subdivides a sphere in uniform flat polygonal faces starting from one of the Platonic solids as the icosahedron or the dodecahedron. One simple technique that increases the number of faces and vertexes by maintaining the same characteristic of the original solid consists in bisecting each face of the icosahedron and projecting the resulting intersection points. The number of vertexes are clearly limited to specific values as 12, 42, 162 or even 252. Intermediate number of points can be obtained using truncated versions in which their surface is composed of multiple different flat regular polygons created by cutting off the vertexes of the solid. As an example, by cutting off all the edges of an icosahedron one can raise the number of vertexes from 12 to 60. Successively, by applying the method described above it is possible to create a new geodesic sphere with 92 vertexes known as geodesic truncated icosahedron.

Choosing the optimal number of vertexes for a LBP is not a trivial task because different factors are involved as the type of texture, the size of the images and the computational power available. An LBP with many vertexes is able to capture all the details of the texture, however, it also captures the noise and it requires a great deal of computation. Instead, with too few vertexes the LBP fails to capture enough details, and hence to reproduce the local texture. Ojala et al. (2002) propose a set of local binary patterns with 8, 16 and 24 pixels which (following the framework proposed by Paulhac et al. (2008)) correspond to 26, 114 and 266 voxels in three dimensions respectively. Therefore, in the remainder of this paper, we use geodesic spheres (derived from the icosahedron) with 42, 92 and 252 vertexes for our LBPs since these values are the closest options available in order to have comparable results.

### 2.2. Extended three-dimensional local binary patterns

Our three-dimensional rotation invariant local binary patterns are a combination of the existing intensity-based



features, difference-based features (Liu et al., 2012) and the three-dimensional descriptor proposed by Paulhac et al. (2008) based on the region-growing algorithm.

Extending the circular local binary pattern developed by Ojala et al. (2002) from two to three dimensions by using the same approach is far from being a simple task. In the classic 2D non-rotation invariant LBP, the texture is simply encoded in a binary number by sampling a geometrical shape such as a circle and by weighting the resulting pixels. In three-dimensions, even if the sampling rate is comparable, the number of points/voxels involved increases dramatically leading us to cope with large integer numbers. Thus, the development of a rotation invariant descriptor is inevitable in three-dimensions due to the intractable number of possible combinations a non-rotation invariant descriptor has ( $2^{42}, 2^{92}, 2^{252}$ ) (Paulhac et al., 2008).

After the sampling procedure of the 2D LBP, the rotation invariance is achieved by applying a circular bit shift. In three dimensions, this sequence of operation is hard to obtain, however, Paulhac et al. (2008) have proposed to bypass this problem by searching directly the uniform conglomerates on the surface of the sphere instead of encoding the texture. A region growing algorithm groups adjacent sampled voxels in order to form regions. If the number of regions is greater than  $v$ , the LBP is defined as "nonuniform" otherwise "uniform". This spherical LBP is intrinsically rotation invariant since the positions of the conglomerates w.r.t the sphere coordinate system does not affect the final result; only the number of regions and the sum of the active sampled voxels determines the LBP code.

Liu et al. (2012) propose a novel set of highly discriminant features based on the LBP method. The combination of the intensity-based and difference-based descriptors proposed in their paper demonstrate superiority in all the experiments against the classic *LBP* and *VAR* operator proposed by Ojala et al. (2002). The intensity-based feature *NI-LBP*, was primarily created to address the deficiencies of the pair *LBP/VAR* in distinguishing particular texture patterns and to increase the overall noise rejection rate. In their theoretical contribution, the *NI-LBP* is claimed to be more discriminant than the classic *LBP* due its higher discriminatory abilities for textures, however, it is important to note that the classic *LBP* in their experiments achieve better results than the single *NI-LBP*. The crucial point in the method proposed by Liu et al. (2012) is in the combination of non-redundant and complementary features as the joint combination of *NI-LBP* and *RD-LBP* descriptors. The combination of these two descriptors outperformed the pair *LBP/VAR* proposed by Ojala et al. (2002) significantly. Ultimately, producing highly discriminant, uncorrelated and complementary descriptors is the key to solve complex problems rather than employing a single method for the solution.

By combining the valuable approaches presented in (Liu et al., 2012) and Paulhac et al. (2008) we define the contrast information *CI-LBP* in three dimensions as:

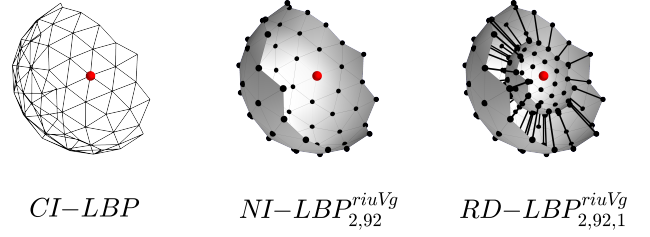


Figure 4: Graphical representation of our extended three-dimensional local binary patterns.

$$CI-LBP = s(g_{(0,0,0)} - \mu_r) \quad (10)$$

where  $\mu_r$  is the grey scale mean of the whole image and  $g_{(0,0,0)}$  is the grey scale of the central voxel. This descriptor aims to highlight intensity variations by considering the image as a whole.

In the same manner we define the three-dimensional intensity-based descriptor *NI-LBP* as:

$$NI-LBP_{R,P}^{riuVg} = \begin{cases} \sum_{p=0}^{P-1} s(g_{(R,p)} - \mu) & \text{if } U(NI-LBP_{R,P}) \leq V \\ P+1 & \text{otherwise} \end{cases} \quad (11)$$

where  $\mu = \frac{1}{P} \sum_{p=0}^{P-1} g_{(R,p)}$  is the grey scale mean of the interpolated neighbouring voxels, the superscript  $g$  indicates the use of a geodesic structure and  $P$  defines the number of voxels in the neighbourhood set. This descriptor takes advantage of the mean operator of the neighbouring voxels to maintain the intensity invariance, to mitigate the effects of intensity variations and therefore to improve the overall noise rejection rate. "LBP thresholding at the value of the central pixel  $g_{(0,0,0)}$  tends to be sensitive to noise, particularly in near-uniform image regions, and smooths weak illumination gradients" (Liu et al., 2012).

Analogously, we define the three-dimensional radial difference descriptor *RD-LBP* as:

$$RD-LBP_{R,P,\delta}^{riuVg} = \begin{cases} \sum_{p=0}^{P-1} s(\Delta_{\delta}^{Rad}) & \text{if } U(RD-LBP_{R,P,\delta}) \leq V \\ P+1 & \text{otherwise} \end{cases} \quad (12)$$

where  $\Delta_{\delta}^{Rad} = g_{(R,p)} - g_{(R-\delta,p)}$ ,  $\delta$  is an integer defining the radial displacement between the external sphere and the internal sphere and the superscript  $g$  indicates the use of a geodesic structure. For instance,  $RD-LBP_{3,252,1}^{riu3g}$  (or  $RD-LBP_{3,252}^{riu3g}$  for brevity) computes the radial difference between a sphere of ( $R = 3, P = 252$ ) and ( $R = 2, P = 252$ ). Although this descriptor appears to be different from the LBP methodology we have seen so far, it becomes in fact equivalent to the LBP proposed by Paulhac et al. (2008) when the radial displacement  $\delta$  approaches the radius of the sphere ( $RD-LBP_{R,P,\delta}^{riu3g} \equiv LBP_{R,P}^{riu3g}$ ). The response to texture variations of an *RD-LBP* whose inner sphere has a radius greater than one will be different from LBP descriptors with one single central voxel. The value of the descriptor *RD-LBP* varies as soon as a change in the texture approaches the outer sphere. This is different from LBPs such as the  $LBP_{R,P}^{riu3g}$  (Paulhac et al., 2008) whose value changes only if the central voxel varies. Figure 4 shows a representation of our extended local binary patterns at  $R = 2$  and  $P = 92$  in three-dimensions.



Table 1: Table of abbreviations

Complete name	Abbreviation	Comment
$NI - LBP_{R,P}^{riuV} / RD - LBP_{R,P',1}^{riuV} / CI - LBP$	$NI / RD / CI - LBP_{R,P'}^{riuV}$	<b>Our proposed method:</b> Joint combination of three extended 3D LBPs at (R,P') spatial resolution where the structure is a series of concentric circles at different heights.
$NI - LBP_{R,P}^{riuVg} / RD - LBP_{R,P,1}^{riuVg} / CI - LBP$	$NI / RD / CI - LBP_{R,P}^{riuVg}$	<b>Our proposed method:</b> Joint combination of three extended 3D LBPs at (R,P) spatial resolution where the frame is a geodesic sphere.
$NI - LBP_{R,P}^{riuVg} + RD - LBP_{R,P,1}^{riuVg} + CI - LBP$	$NI + RD + CI - LBP_{R,P}^{riuVg}$	<b>Our proposed method:</b> Concatenation of three extended 3D LBPs at (R,P) spatial resolution where the frame is a geodesic sphere.
$LBP_{R,P}^{riu3}$	-	<b>(Paulhac et al., 2008):</b> Three-dimensional fully rotation invariant LBP based on the region growing algorithm.
$kLBP_{R,P}^{riu3D}$	-	<b>(Banerjee et al., 2013):</b> Three-dimensional fully rotation invariant LBP based on the norm of the spherical harmonics frequency components and the kurtosis.
$rLBP_{R,P}^n$	-	<b>(Fehr and Burkhardt, 2008):</b> Three-dimensional fully rotation invariant LBP based on the spherical harmonics and the search of the minimum correlation over all angles.
$VLBP_{L,R,P}^{u2}$	-	<b>(Zhao and Pietikäinen, 2007a):</b> Three-dimensional single axe rotation invariant LBP in the form of spiral. The LBP employed is the 2D uniform not rotation invariant LBP.
$VLBP_{L,R,P}^{riu2}$	-	<b>(Zhao and Pietikäinen, 2007a):</b> Three-dimensional single axe rotation invariant LBP in the form of spiral. The LBP employed is the 2D uniform and rotation invariant LBP.
$LBP - TOP_{P_{XY}, P_{XT}, P_{YT}, R_X, R_Y, R_T}^{u2}$	$LBP - TOP_{R,P}$	<b>(Zhao and Pietikäinen, 2007a):</b> Three-dimensional rotation invariant LBP composed of 3 classic 2D LBP arranged in an orthogonal fashion.
$vcLBP_{R,P}$	-	<b>(Liu et al., 2011):</b> Three-dimensional fully rotation invariant LBP based on unique rotation, reflection and translation invariant patterns (congruent patterns).
$NI - LBP_{R,P}^{riu2} / RD - LBP_{R,P,1}^{riu2} / CI - LBP$	$NI / RD / CI - LBP_{R,P'}^{riu2}$	<b>(Liu et al., 2012):</b> Joint combination of three 2D rotation invariant extended LBPs.

### 3. Results and discussion

In this section, we evaluate the performance of our proposed descriptors by using a single and multiple spatial resolutions with different number of regions  $V$ . We perform classification tasks against state-of-the-art two-dimensional and three-dimensional LBP descriptors on a dataset of synthetic textures in order to firstly demonstrate the utility of the third dimension and secondly to demonstrate the discriminatory power of our improved descriptors. The first descriptor considered here is the 2D joint LBP  $NI/RD/CI - LBP_{R,P}^{riu2}$  proposed by Liu et al. (2012) which produce excellent results in classifying two-dimensional datasets. The first 3D descriptors under consideration are the volume LBP  $VLBP_{L,R,P}^{u2}$ , the  $VLBP_{L,R,P}^{riu2}$  and the  $LBP - TOP_{P_{XY}, P_{XT}, P_{YT}, R_X, R_Y, R_T}^{u2}$  proposed by Zhao and Pietikäinen (2007a). The first 3D fully rotation invariant LBP employed here is the  $LBP_{R,P}^{riu3}$  proposed by Paulhac et al. (2008) followed by a couple of variations of the 3D  $rLBP_{R,P}^n$  (Fehr and Burkhardt, 2008) based on the spherical harmonics and the  $kLBP_{R,P}^{riu3D}$  proposed by Banerjee et al. (2013) based as well on the spherical harmonics. Another interesting method we examine here is the 3D fully rotation invariant volumetric congruent LBP  $vcLBP_{R,P}$  proposed by Liu et al. (2011).

#### 3.1. Dataset

The dataset employed here is a set of three-dimensional synthetic greyscale texture images constructed from two-dimensional textures like Brodatz, fractal textures and etc. This dataset has been created by Paulhac et al. (2009) and is freely available<sup>1</sup>. The full database offers 92 classes composed of ten

images constructed by using different synthesis methods such as interpolation, Fourier Transform, geometrical shape insertion and a combination of these methods. For each method, various manipulations have been applied as the rotation, addition of noise, sub-sampling and smoothing. The rotation has been randomly applied on the three axes. For all of our tests, we use the interpolated dataset (30 classes) composed of images of  $64 \times 64 \times 64$  voxels and in particular the rotated version in order to evaluate the rotation invariance property of our algorithm. Figure 5 shows some examples of 3D textures from the 30 texture classes composing the interpolated dataset (not rotated).

#### 3.2. Comparing textures

Defining the local binary pattern structure and how to calculate the resulting LBP code are the first steps toward the comparison of two textures. We define here the term "converting an image" as the process that extracts an LBP code for each pixel in an image; the result is also an image where each pixel/voxel ranges from 0 to  $P - 1$ . Hence, a simple method to compare two textures consists in comparing the fixed size histograms derived from the converted images. In addition, joint combination and/or concatenation of histograms of different LBPs are proven to be effective (Liu et al. (2012)).

#### 3.3. Variable selection

An important task in the implementation of an algorithm is the selection of its coefficients. The number of voxels  $P$  on the LBP sphere is an aspect to consider, but it is not crucial. Too few points could lead to the loss of information, on the other hand, too many points increase the complexity and the length of the resulting histograms. The number of regions  $V$  is one

<sup>1</sup>[http://www.rfai.li.univ-tours.fr/PublicData/3D\\_Textures/3Dsynthetic\\_images\\_database.html](http://www.rfai.li.univ-tours.fr/PublicData/3D_Textures/3Dsynthetic_images_database.html) (12.02.2017)

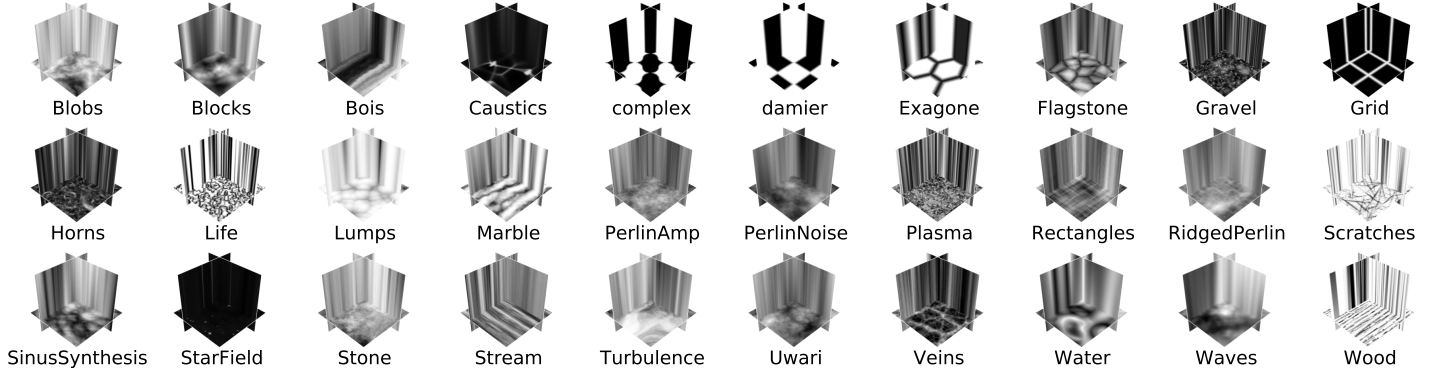


Figure 5: Example images from the three-dimensional dataset of synthetic textures (Paulhac et al., 2009): 30 classes, interpolated and not rotated.

of the parameters whose effects on the algorithm need to be considered.

- For the volume local binary pattern  $VLBP_{L,R,P}$ , we decide to stack five 2D LBP composed of 16 voxels each by forming a spiral. The resulting descriptor is ended up being formed by 82 voxels.
- For the descriptor  $LBP - TOP_{R,P}$ , we employ three classic 2D  $LBP_{2,16}^{pi2}$  arranged orthogonally.
- Concerning the 3D descriptor  $kLBP_{R,P}^{ri3D}$  exploiting the spherical harmonics, the number of bands  $n$ , the number of bins and the range for each histogram must be selected. We choose the number of bands  $n$  equal to six as advised in Banerjee et al. (2013) whereas for the number of bins we perform several experiment to select the appropriate one. The results are shown in Table 2 where the number of bins refers to the concatenation of  $n + 1$  variables. Regarding the radius and the sampling rate, Banerjee et al. (2013) state that "the increase in sampling rate, improves the spherical function approximation, which translates to improved discriminative ability of the descriptor". To verify this statement, we perform experiments at different spatial resolutions. The range of the  $n + 1$  histograms has been empirically estimated.
- For the second descriptor based on the spherical harmonics  $rLBP_{R,P}^n$  (Fehr and Burkhardt, 2008), we perform several experiments at different spatial resolutions and with two different spherical harmonics expansion bands, six ( $n = 6$ ) and eleven ( $n = 11$ ) to verify if increasing the SH resolution helps improving its discriminatory power. Experiments at the spatial resolution of (2,252) and (3,252) are not conducted due to numeric instability. Also here, for the number of bins we perform several experiments to select the appropriate one. The results are shown in Table 2. The range of the histogram has been empirically estimated.
- Concerning the volumetric congruent  $vcLBP_{R,P}$  (Liu et al., 2011), the number of voxels forming the sphere and the radius must be selected. To implement this method, an eigendecomposition of a symmetric distance-preserving

matrix is required to be performed for each pattern ( $2^P$ ) in order to identify all possible congruent geometries. Therefore, admitting a roughly constant amount of time to perform the eigendecomposition, with the increase of the number of voxels in the spherical frame the time required to pre-compute all the patterns increases exponentially becoming an intractable method for a regular desktop machine. This therefore makes us use a geodesic (octahedron) sphere composed of only 18 vertexes. Due to the small number of vertexes, we also decide to perform two distinct experiments using a radius of two ( $R = 2$ ) and a radius of one ( $R = 1$ ). In our implementation, after the clustering process, we identify 6'391 different groups of congruent shapes.

### 3.4. Evaluation method

We aim to compare the classification accuracies of our descriptors using 30 different texture classes where each class is composed of ten randomly rotated images. We convert the images in the dataset using the LBP algorithm, successively, we extract LBP histograms by considering the entire image. The K-nearest-neighbours ( $k = 1$ ) classifier was used with the L1 metric to emphasize the discriminatory power of each descriptor. The resulting accuracy is an average over 100 runs where we randomly take four out of ten texture images in each class for training and the remaining six for testing. We have not performed any normalization of the features.

### 3.5. Experiment #1

The first experiment aims to compare the classification accuracies of our algorithm against state-of-the-art 2D and 3D descriptors (Figure 6).

We compare the performance of our algorithm with those of the two-dimensional joint LBP  $NI/RD/CI - LBP_{2,16}^{riu2}$  proposed by Liu et al. (2012), the 3D rotation invariant  $LBP_{2,98}^{riu3}$  proposed by Paulhac et al. (2008), the volume LBP  $VLBP_{1,16,2}^{pi2}$  and  $VLBP_{1,16,2}^{riu2}$  as well as the 3D  $LBP - TOP_{2,16}$  proposed by Zhao and Pietikäinen (2007a). Moreover, we present the accuracies of different 3D  $kLBP_{R,P}^{ri3D}$  algorithms proposed by Banerjee et al. (2013) as well as different  $rLBP^n$  algorithms. We also present the results of the 3D volumetric congruent  $vcLBP_{2,18}$  and

Table 2: Classification accuracies [%] for the methods proposed by Banerjee et al. (2013) and Fehr and Burkhardt (2008) at different spatial resolutions and number of bins in the final histogram. For the method  $rLBP_{R,P}^n$  (Fehr and Burkhardt, 2008) we also analyse the effect of two different number of bands in the spherical harmonics expansion,  $n = 6$  and  $n = 11$ .

$kLBP_{\{\downarrow\}}^{riu3D}$	Number of histogram bins								
	35	70	350	700	3500	7000	35000	70000	105000
(1,42)	86.2	<b>87.9</b>	86.9	85.4	81.2	79.2	78.0	76.8	76.7
(1,92)	87.5	89.8	<b>90.3</b>	89.5	87.1	86.4	82.7	80.4	79.2
(2,42)	91.6	<b>93.5</b>	92.9	91.7	87.6	85.7	82.8	82.7	82.6
(2,92)	91.7	94.2	<b>94.6</b>	<b>94.6</b>	93.5	92.4	89.4	87.2	86.9
(2,252)	91.0	93.6	<b>94.8</b>	94.6	94.6	94.3	92.4	91.9	92.4
(3,92)	88.3	92.5	94.2	<b>94.5</b>	93.9	93.3	90.7	89.5	88.9
(3,252)	87.7	92.2	94.6	<b>94.9</b>	<b>94.9</b>	94.4	94.0	93.0	92.4
$rLBP_{\{\downarrow\}}^{n=6}$	Number of histogram bins								
	5	10	50	100	500	1000	5000	10000	15000
(1,42)	54.8	59.5	57.7	56.8	58.0	59.3	<b>61.8</b>	61.6	60.5
(1,92)	58.2	59.9	63.8	63.3	69.9	72.9	<b>77.4</b>	76.0	74.3
(2,42)	50.8	53.1	56.6	57.5	62.8	64.7	<b>66.5</b>	64.4	65.0
(2,92)	52.5	56.1	57.9	58.4	68.4	73.9	81.7	83.2	<b>83.4</b>
(2,252)	-	-	-	-	-	-	-	-	-
(3,92)	49.4	54.3	58.3	59.7	66.9	70.0	79.3	<b>79.7</b>	79.3
(3,252)	-	-	-	-	-	-	-	-	-
$rLBP_{\{\downarrow\}}^{n=11}$	Number of histogram bins								
	5	10	50	100	500	1000	5000	10000	15000
(2,42)	3.3	8.4	57.5	<b>74.3</b>	70.6	67.2	53.8	46.4	43.0
(2,92)	44.4	48.2	67.2	77.7	<b>85.5</b>	83.2	75.0	70.6	67.5
(2,252)	-	-	-	-	-	-	-	-	-

Table 3: Summary of the best classification accuracies [%] reported in Table 2 and for our proposition of extended 3D geodesic LBP at different spatial resolutions.

Descriptor	Spatial resolution						
	(1,42)	(1,92)	(2,42)	(2,92)	(2,252)	(3,92)	(3,252)
$kLBP^{riu3D}$	87.9	90.3	93.5	94.8	94.6	94.5	<b>94.9</b>
$rLBP^{n=6}$	61.8	77.4	66.5	<b>83.4</b>	-	79.7	-
$rLBP^{n=11}$	-	-	74.3	<b>85.5</b>	-	-	-
$NI/RD/CI - LBP^{riu3g}$	94.2	<b>95.3</b>	92.9	93.4	93.7	89.4	88.1

$vcLBP_{1,18}$  algorithms developed by Liu et al. (2011). Regarding the 2D  $NI/RD/CI - LBP_{2,16}^{riu2}$ , the three-dimensional texture images are converted slice by slice.

Table 2 shows an extensive comparison of results at different spatial resolutions and number of bins for the two methods  $rLBP^n$  and  $kLBP^{riu3D}$  based on the spherical harmonics (Fehr and Burkhardt, 2008; Banerjee et al., 2013). For these two methods, the number of bins in the resulting histograms plays an important role, instead, classic LBP methods are defined to provide fixed length feature vectors.

For the descriptor proposed by Banerjee et al. (2013) ( $kLBP^{riu3D}$ ), an inaccurate selection of the number of bins could lead to a loss in the accuracy score up to 15 % whereas for the descriptor proposed by Fehr and Burkhardt (2008) ( $rLBP^n$ ) up to even 40 % on this dataset. With a radius of two and three the descriptor  $kLBP^{riu3D}$  produces very high scores of 94.8 % and 94.9 % at the spatial resolution of (2,252) and (3,252) respectively.

On the other hand, with a spherical radius of one ((1,42) and (1,92)), the method  $kLBP^{riu3D}$  appears to be slightly less capable of capturing high significant textural details compared to larger radii. The spherical harmonics frequency components forming the resulting feature vector in this case are not discriminant enough for reasons that we are unable to explain and that goes behind the scope of this paper. In addition, it can be noted that the increase in sampling rate, for example from (2,42) to (2,252), slightly improves the results as stated in (Banerjee et al., 2013) but to the detriment of the execution speed. Table 4 shows how the execution time for the method  $kLBP^{riu3D}$  increases dramatically with the increase of the sampling rate. The amount of time required to convert an image with the descriptor  $kLBP_{x,42}^{riu3D}$  is more than 20 times the amount required using the descriptor  $kLBP_{x,252}^{riu3D}$ , regardless the radius of the LBP sphere.

Regarding the descriptor proposed by Fehr and Burkhardt (2008) ( $rLBP^n$ ), the results obtained are not as impressive as the scores produced by its counterpart  $kLBP^{riu3D}$  but it still among the most discriminant method analysed in this paper. Opposite to the method  $kLBP^{riu3D}$ , the descriptor  $rLBP^n$  appears to be less affected by decreases in the score due to the radius but more conditioned by the number of voxels in the spherical frame. The increase in the sampling rate from (2,42) to (2,92) clearly improves the results, however, it also increases the instability of this method due to numeric approximations and operations on large numbers. Because of this issue, the results at the spatial resolution of (2,252) and (3,252) are not provided here. Focusing on the comparison of the descriptors with expansion bands  $n = 6$  and  $n = 11$ , the latter produces a slightly better result of 85.5 % at the spatial resolution of (2,92) in contrast with the 83.4 % accuracy of the  $rLBP^{n=6}$  at the same resolution. Thus, the increase of the number of bands produces some improvements, but also in this case, to the detriment of the execution time. The time required to execute the descriptor with expansion band eleven ( $rLBP^{n=11}$ ) is seven times the amount required by the descriptor employing a number of bands equal to six ( $rLBP^{n=6}$ ). It is important to note that the vast majority of the delay introduced by this method is due to the size of the data to be converted from the frequency domain to time domain using the inverse Fourier transform (IFFT).

Table 3 summarises the classification accuracies obtained in Table 2 and in addition with the results of our proposed joint method  $NI/RD/CI - LBP^{riu3g}$  at different spatial resolutions. For the descriptor proposed by Banerjee et al. (2013) ( $kLBP^{riu3D}$ ) it is interesting to note how the accuracy scores become constants from the spatial resolution (2,42) up to (3,252). On the other hand, our proposition  $NI/RD/CI - LBP^{riu3g}$  produces relatively constant results from (1,42) to (2,252) with a peak at spatial resolution (1,92).

In addition to the extensive comparison presented so far on the descriptor proposed by Banerjee et al. (2013), Fehr and Burkhardt (2008), we present in Figure 6 an overall view of the accuracy scores of all methods examined.

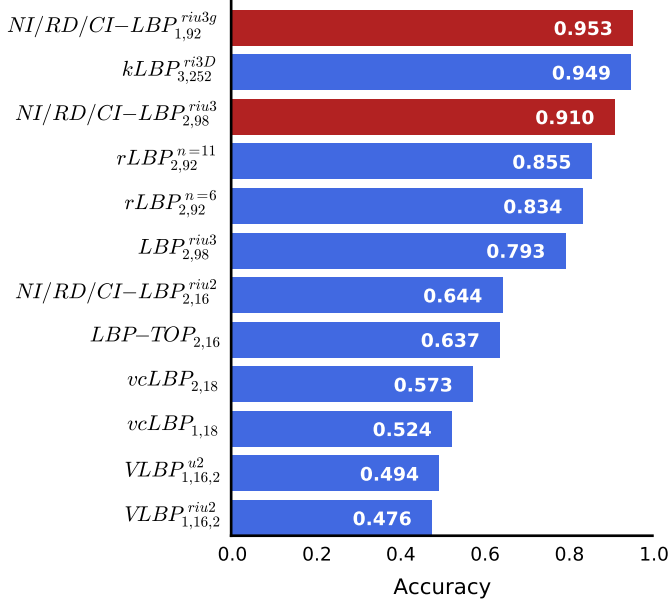


Figure 6: Summary and comparison of accuracy scores obtained from different state-of-the-art 2D and 3D hand-crafted texture descriptors examined in this paper. The bars in red represent our accuracy results and the blue bars correspond to the results from the other methods examined.

The first main fact to note in Figure 6 is the preeminence of the three-dimensional fully rotation invariant descriptors. These methods perform in general well on this dataset rather than two-dimensional and partially rotation invariant descriptors. This behaviour is not a surprise since we aim to highlight the discriminatory power and the rotation invariant property of the descriptors using a dataset of randomly rotated textures. The only descriptors against the flow are the fully rotation invariant volumetric congruent 3D LBP  $vcLBP_{1,18}$  and  $vcLBP_{2,18}$  whose accuracies are as low as 52.4 % and 57.3 % respectively. The low sampling rate for this descriptor plays an important role on the final classification results whereas the proportion between sampling rate and radius does not. The choice of using such small number of voxels in the spherical frame is inevitable because of the computational costs of calculating all possible congruent patterns in advance. To mitigate this constraint, we decided to reduce the radius in order to adapt the proportion between radius and sampling rate. However, such a reduction of the radius has not improved the final result for  $vcLBP_{1,18}$ .

The two-dimensional LBP  $NI/RD/CI - LBP_{2,16}^{riu2}$  presents a moderate 64.4 % of accuracy in classifying 30 different classes outperforming the volume LBPs  $VLBP_{1,16,2}^{u2}$  and  $VLBP_{1,16,2}^{riu2}$  which they produce 49.4 % and 47.6 % accuracies respectively. These two methods ( $VLBP$  and  $NI/RD/CI - LBP_{2,16}^{riu2}$ ), share common characteristics such as the same underlying LBP descriptor ( $LBP_{2,16}^{riu2}$ ) and the rotation invariance capability along one direction only despite the fact that one can be considered a 3D LBP. These common characteristics bring the two descriptors extracting the same information from the texture, however, the 2D LBP  $NI/RD/CI - LBP_{2,16}^{riu2}$  produces better results due to the joint combination of three complementary descriptors.

The last LBP that presents a moderate score is the 3D  $LBP - TOP_{2,16}^{u2}$  which combines three classic LBP in an orthogonal fashion. This descriptor reports 63.7 % accuracy similar to the 2D  $NI/RD/CI - LBP_{2,16}^{riu2}$  which as well is formed by a combination of three classic 2D LBP descriptors.

Among the other three-dimensional fully rotation invariant descriptors, the LBP based on the region growing algorithm  $LBP_{2,98}^{riu3}$  demonstrates a good 79.3 % accuracy similar to the results presented in their paper (Paulhac et al., 2008).

The 3D fully rotation invariant method proposed by Banerjee et al. (2013) ( $kLBP^{riu3D}$ ) produces an impressive result similar to our method, 94.9% at the spatial resolution (3,252) with a much smaller histogram of 700 elements in contrast with the 17672 bins histogram produced by our joint 3D proposition  $NI/RD/CI - LBP_{x,92}^{riu3g}$ . Moreover, the increase in sampling rate for the method  $kLBP^{riu3D}$  improves the classification accuracy as shown in Table 2 but with a significant increase in execution time.

The second 3D fully rotation invariant descriptor analysed here (Fehr and Burkhardt, 2008) with the number of bands equal to six  $rLBP_{2,92}^{n=6}$  also demonstrates a good 83.4 % accuracy using a histogram of 15000 bins. The attempt to improve the discriminatory power by increasing the spherical approximation from six to eleven ( $rLBP_{2,92}^{n=11}$ ) produces a slightly better score of 85.5 % accuracy of correct classified textures at the same spatial resolution but by employing a much smaller histogram of 3500 elements. Thus, for  $rLBP_{2,92}^{n=6}$  proposed by Fehr and Burkhardt (2008), increasing the number of bands from six to eleven slightly improves the final classification accuracy however, also here, the execution time increases dramatically from 56 seconds to 377 seconds on our machine (Table 4).

Focusing on the results of our proposed methods, the joint descriptor  $NI/RD/CI - LBP_{2,98}^{riu3}$  and the joint geodesic  $NI/RD/CI - LBP_{1,92}^{riu3g}$  produce interesting results of **91.0 %** and **95.3 %** accuracies respectively standing among the top state-of-the-art hand-crafted three-dimensional texture descriptors. The intensity-based features as well as the difference-based features successfully increase the performance of the simple 3D  $LBP_{2,98}^{riu3}$  proposed by Paulhac et al. (2008) by almost 12 % and the geodesic structure also contributes by a good 2 % on this dataset.

Besides the classification accuracies, we examine the execution times. Table 4 shows the time required to convert all the voxels of an image of  $64 \times 64 \times 64$ . It is noted that, the time required to jointly combine or concatenate histograms is equivalent. The heart of the algorithms proposed here are coded in C and wrapped in Python. Each experiment is performed on a single thread of a quad core (hyper-threading) i7-4700MQ 2.4GHz 64 bits Intel processor.

The number of points in the spherical frame is the main cause for a long execution time for our descriptors due to the numerous voxel interpolations and the region growing algorithm.

Table 4: Time required to convert an image composed of  $64 \times 64 \times 64$  voxels into an LBP histogram using a single thread of a quad core (hyper-threading) i7-4700MQ 2.4GHz 64 bits Intel processor.

Descriptor	Execution time [s]
$kLBP^{riu3D}_{x,252}$	460
$rLBP^{n=11}_{x,92}$	377
$rLBP^{n=11}_{x,42}$	370
$vcLBP_{x,18}$	160
$NI/RD/CI - LBP^{riu3g}_{x,252}$	104
$kLBP^{riu3D}_{x,92}$	71
$rLBP^{n=6}_{x,92}$	56
$rLBP^{n=6}_{x,42}$	44
$NI/RD/CI - LBP^{riu3}_{x,98}$	48
$NI/RD/CI - LBP^{riu3g}_{x,92}$	21
$LBP^{riu3}_{x,98}$	19
$kLBP^{riu3D}_{x,42}$	19
$NI/RD/CI - LBP^{riu3g}_{x,42}$	8
$VLBP^{riu2}_{1,16,2}$	6
$VLBP^{u2}_{1,16,2}$	1
$LBP - TOP^{u2}_{2,16}$	1

A greater number of points involves a greater searching space for the agglomeration process. For the descriptor proposed by Fehr and Burkhardt (2008) ( $rLBP^n$ ), increasing the spherical harmonic approximation is the major causes of a greater delay not because of the complexity involved but because of the increased number of samples involved in the inverse Fourier transform. On the other hand, the bottleneck for the descriptor proposed by Banerjee et al. (2013) ( $kLBP^{riu3D}$ ) is in the computation of the frequency components norms which their computational cost depends upon the sampling rate. Ultimately, our descriptor  $NI/RD/CI - LBP^{riu3g}_{1,92}$  demonstrates one of the highest accuracy on this specific dataset and a short execution time in comparison with other methods.

### 3.6. Experiment #2

The second experiment aims to compare the performance of our descriptors at different spatial resolutions and with different number of regions  $v$  (Table 5). Jointly combining individual local binary patterns has been proven to be effective (Ojala et al., 2002), however, the time required to process joint histograms increases with the increase of the number of voxels in the neighbourhood. A simple concatenation produces shorter histograms that may be more beneficial in some applications. For this reason, we perform analysis on individual LBP as well as joint combinations (/) and concatenations (+).

The concatenation or the joint combination of descriptors produces higher classification scores than individual descriptors. The intensity-based descriptor  $NI - LBP$  is the most discriminant among the individual LBPs followed by the difference-based  $RD - LBP$ . Unexpectedly, the concatenation  $NI + RD$  at the spatial resolution of (1,92) produces the best scores of **95.7 %** and **95.4 %** with  $v = 3$  and  $v = 4$  respectively. This result is in contrast with the assumption that jointly

combining LBP histograms produces better results as reported in Ojala et al. (2002). Adding the contrast information  $CI$  to the  $NI/RD$  and the  $NI + RD$  descriptors has not improved the classification scores as for Liu et al. (2012). This discrepancy may be attributable to the noise or to the characteristics of the dataset employed here. Increasing the number of regions  $v$  from three to four produces similar results with sporadic faint improved results in both sides. Therefore, according to the results, there is no evident benefit in using a larger number of regions  $v$ .

Figure 7 shows the percentage of uniform patterns for the intensity-based descriptors  $NI - LBP$  and the difference-based  $RD - LBP$  with  $v = 3$  and  $v = 4$  for each class in the dataset. Due to its nature, the proportion of uniform pattern for the  $NI - LBP$  is overall greater than the  $RD - LBP$ . In both descriptors, a very low percentage in accuracy is reported in some classes mostly due to the type of the texture. For instance, the classes Life and Plasma are two vibrant textures that produce multiple regions on the LBP spheres and consequently more "nonuniform" LBPs. Increasing the radius clearly decreases the percentage of uniform patterns. A low percentage of uniform patterns could lead to misclassifications due to the scarcity of the information held by the histograms, however, this is not a general rule as we can see in Table 6. The misclassification rates for the classes Plasma and Life are in line with the other scores whereas other classes such as Turbulence and Uwari have a high percentage of uniform patterns but with relatively high misclassification rates.

At the spatial resolution of (3,252) (blue curves), the average percentage of uniform codes is considerably lower in comparison with the other resolutions due to the scale of the texture/motif. The size of the LBP compared to the scale of the texture is clearly an important factor. A very dense and vibrant texture can be better analysed using an LBP with a small radius rather than a large one. The general reduction of uniform pattern has the impact of decreasing the overall discriminability of the descriptor which is probably the reason of the decrease in the classification accuracies reported in Table 5 for the descriptors with radius three. Moreover, for this specific case, increasing the number of regions  $v$  results in a slightly greater percentage of uniform patterns but with no relevant improvements in the classification scores.

### 3.7. Experiment #3

In the third experiment, we perform a multi-resolution analysis. Table 7 shows the accuracies in classifying the 30 different classes by using some combinations of descriptor at different resolutions.

In the same way as in Experiment #2, the concatenated LBP  $NI + RD$  at (1,42)+(2,92) with  $v = 4$  presents the best score of **95.9 %**. The joint  $NI/RD$  at (1,42)+(2,92) follows with **95.5 %** and **95.6 %** accuracies with  $v = 3$  and  $v = 4$  respectively. Overall, the combination of multiple descriptors at different spatial resolutions leads to better results but by a small amount. The descriptors at (1,42) and (2,92) demonstrate very good classification accuracies whereas the descriptors at (3,252) tends to

Table 5: Classification accuracies [%] for different individual LBP, joint and concatenated LBP with different number of regions V

$\{\downarrow\}-LBP_{\{\rightarrow\}}^{riuVg}$	V=3							V=4						
	(1,42)	(1,92)	(2,42)	(2,92)	(2,252)	(3,92)	(3,252)	(1,42)	(1,92)	(2,42)	(2,92)	(2,252)	(3,92)	(3,252)
<i>NI</i>	89.3	91.0	89.2	90.6	91.0	84.0	84.4	87.5	89.3	86.3	88.1	89.7	82.4	84.2
<i>RD</i>	82.6	85.3	85.2	84.5	79.0	76.9	67.4	82.7	85.2	87.9	88.8	85.0	84.0	71.8
<i>CI</i>	22.3	22.3	20.6	20.6	20.6	20.6	20.6	22.3	22.3	20.6	20.6	20.6	20.6	20.6
<i>NI/RD</i>	<b>94.3</b>	<b>95.4</b>	92.9	93.5	93.4	87.5	86.6	<b>94.4</b>	<b>95.3</b>	93.8	93.9	94.0	89.2	89.3
<i>NI/CI</i>	91.0	92.5	89.4	90.5	91.5	85.9	87.4	90.0	91.7	86.4	88.3	89.8	83.3	86.1
<i>RD/CI</i>	84.7	86.1	87.2	85.4	82.1	78.3	68.0	85.1	86.0	88.3	89.4	85.9	84.8	75.9
<i>NI/RD/CI</i>	<b>94.2</b>	<b>95.3</b>	92.9	93.4	93.7	89.4	88.1	<b>94.2</b>	<b>95.2</b>	93.2	93.7	93.8	89.3	89.5
<i>NI + RD</i>	<b>94.6</b>	<b>95.7</b>	92.7	92.8	92.6	87.7	86.2	<b>94.9</b>	<b>95.4</b>	93.7	93.7	93.9	89.1	87.4
<i>NI + CI</i>	89.3	91.1	89.0	89.8	90.5	85.2	85.4	88.0	90.0	86.2	87.5	88.6	82.6	85.2
<i>RD + CI</i>	82.0	83.2	85.0	83.7	79.2	76.2	64.3	82.3	83.0	86.0	86.9	84.0	82.0	72.2
<i>NI + RD + CI</i>	<b>94.2</b>	<b>95.1</b>	92.4	93.0	92.9	86.8	87.4	<b>94.3</b>	<b>95.2</b>	92.2	92.6	92.7	88.7	87.5

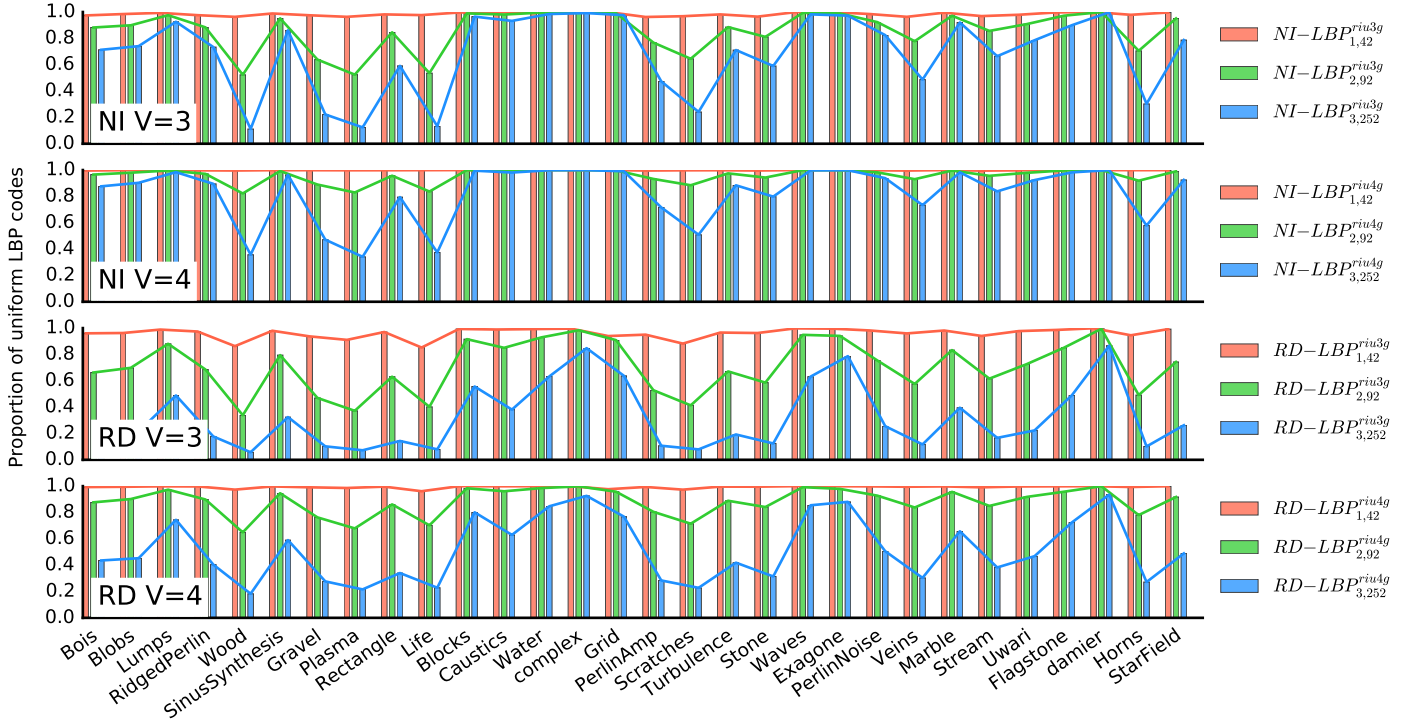


Figure 7: Proportion of uniform LBP codes per class for different descriptors at different spatial resolution and number of region V

 Table 6: Misclassification rates [%] per class for different individual LBPs at different spatial resolutions with  $V = 3$  (cells greater than 30 % are highlighted)

	Bois	Blobs	Lumps	RidgedPerlin	Wood	SinusSynthesis	Gravel	Plasma	Rectangle	Life	Blocks	Caustics	Water	complex	Grid	PerlinAmp	Scratches	Turbulence	Stone	Waves	Exagone	PerlinNoise	Veins	Marble	Stream	Uwar	Flagstone	damier	Horns	Starfield	Average
<i>NI - LBP^{riu3g}_{(1,42)}</i>	6.7	0.0	<b>30.2</b>	26.8	20.0	0.0	6.8	26.7	0.0	0.2	1.0	0.0	13.3	0.0	3.5	0.0	18.0	<b>30.3</b>	10.8	28.3	0.0	28.3	3.3	0.2	16.3	<b>40.2</b>	4.3	0.0	0.0	0.0	10.5
<i>RD - LBP^{riu3g}_{(1,42)}</i>	<b>31.0</b>	<b>43.8</b>	<b>32.8</b>	24.5	8.8	19.0	23.0	0.0	7.2	4.8	22.5	0.0	14.7	0.0	2.3	<b>30.7</b>	3.0	<b>45.5</b>	19.3	0.5	0.0	19.2	1.3	<b>35.8</b>	<b>52.5</b>	<b>47.8</b>	17.8	0.0	17.8	0.3	17.5
<i>NI - LBP^{riu3g}_{(2,92)}</i>	24.7	2.7	4.0	20.3	1.8	0.0	0.0	20.8	0.0	12.7	8.3	0.0	<b>34.0</b>	0.0	0.0	0.0	0.0	23.5	16.7	20.3	0.0	30.0	0.2	2.2	10.8	<b>38.5</b>	0.8	0.0	0.0	0.5	9.1
<i>RD - LBP^{riu3g}_{(2,92)}</i>	<b>55.2</b>	<b>30.8</b>	1.8	<b>36.2</b>	20.7	3.8	2.6	15.0	7.5	<b>37.3</b>	1.7	0.0	0.0	0.0	5.5	19.2	<b>50.5</b>	<b>40.3</b>	26.8	0.0	0.0	<b>33.0</b>	5.7	7.5	23.7	27.2	0.2	0.0	12.7	0.0	15.5
<i>NI - LBP^{riu3g}_{(3,252)}</i>	<b>34.5</b>	14.2	12.8	<b>32.2</b>	0.8	0.2	3.2	28.5	0.0	19.8	24.8	0.0	25.5	0.0	0.0	0.0	0.0	<b>46.2</b>	<b>50.5</b>	1.2	0.0	<b>48.0</b>	4.8	8.2	<b>44.7</b>	<b>58.3</b>	0.7	0.0	3.5	0.0	15.4
<i>RD - LBP^{riu3g}_{(3,252)}</i>	<b>92.0</b>	<b>57.3</b>	<b>37.7</b>	22.8	<b>30.7</b>	<b>47.5</b>	19.2	17.2	6.2	<b>61.0</b>	<b>41.0</b>	10.5	<b>35.0</b>	0.0	4.0	<b>69.0</b>	<b>59.0</b>	<b>64.7</b>	<b>46.0</b>	5.0	0.0	<b>37.5</b>	<b>34.5</b>	24.3	<b>37.8</b>	<b>53.3</b>	12.5	0.2	<b>50.5</b>	2.7	32.6



Table 7: Multiresolution accuracies [%] for different combinations of joint LBPs with different number of regions V

Descriptors	Number of bins	V=3	V=4
$NI/RD - LBP^{riuVg}_{(1,42)+(2,92)}$	10772	(44*44)+(94*94)	<b>95.5%</b> <b>95.6%</b>
$NI/RD - LBP^{riuVg}_{(1,42)+(3,252)}$	66452	(44*44)+(254*254)	95.0% 95.4%
$NI/RD - LBP^{riuVg}_{(2,92)+(3,252)}$	73352	(94*94)+(254*254)	92.1% 93.5%
$NI/RD - LBP^{riuVg}_{(1,42)+(2,92)+(3,252)}$	75288	(44*44)+(94*94)+(254*254)	95.2% 95.1%
$NI/RD/CI - LBP^{riuVg}_{(1,42)+(2,92)}$	21544	(44*44*2)+(94*94*2)	<b>94.8%</b> <b>94.7%</b>
$NI/RD/CI - LBP^{riuVg}_{(1,42)+(3,252)}$	132904	(44*44*2)+(254*254*2)	94.7% 94.7%
$NI/RD/CI - LBP^{riuVg}_{(2,92)+(3,252)}$	146704	(94*94*2)+(254*254*2)	92.9% 93.4%
$NI/RD/CI - LBP^{riuVg}_{(1,42)+(2,92)+(3,252)}$	150576	(44*44*2)+(94*94*2)+(254*254*2)	94.8% 94.4%
$NI + RD - LBP^{riuVg}_{(1,42)+(2,92)}$	276	(44+44)+(94+94)	<b>95.3%</b> <b>95.9%</b>
$NI + RD - LBP^{riuVg}_{(1,42)+(3,252)}$	596	(44+44)+(254+254)	94.8% 95.0%
$NI + RD - LBP^{riuVg}_{(2,92)+(3,252)}$	696	(94+94)+(254+254)	91.8% 93.0%
$NI + RD - LBP^{riuVg}_{(1,42)+(2,92)+(3,252)}$	784	(44+44)+(94+94)+(254+254)	95.0% 95.0%
$NI + RD + CI - LBP^{riuVg}_{(1,42)+(2,92)}$	280	(44+44+2)+(94+94+2)	<b>94.2%</b> <b>94.7%</b>
$NI + RD + CI - LBP^{riuVg}_{(1,42)+(3,252)}$	600	(44+44+2)+(254+254+2)	94.2% 94.7%
$NI + RD + CI - LBP^{riuVg}_{(2,92)+(3,252)}$	700	(94+94+2)+(254+254+2)	91.8% 93.4%
$NI + RD + CI - LBP^{riuVg}_{(1,42)+(2,92)+(3,252)}$	790	(44+44+2)+(94+94+2)+(254+254+2)	93.8% 94.4%

decrease the overall performance of the strongest LBPs in this dataset. The major drawback of the multiresolution approach is the size of the resulting histograms. Joint combinations of histograms cause higher numerical costs, however, as we have seen before, the simple concatenation produces satisfying results that may be more beneficial when fast calculations are required. Also here, increasing the number of regions  $v$  from three to four have not produced evident improvements.

### 3.8. Experiment #4

In the fourth experiment we exploit the characteristics of our 3D local binary patterns to solve a real-world problem.

In medical environments, advanced magnetic resonance imaging (MRI) techniques can be used to improve understanding of brain injuries. Susceptibility-weighted imaging (SWI) is an MRI technique highly sensitive to magnetic susceptibility differences between tissues and oxygenated/deoxygenated blood. Such imaging method is suitable for texture analysis methods, however, one main issue of SWI is the level of oxygen which has the effect of intensifying how veins and tissues appear.

In this experiment we aim to distinguish the oxygenated images (patients who are given oxygen during the MRI acquisition) from the non-oxygenated ones using our extended 3D descriptors.

Our clinical dataset collected as a part of a clinical study (NRES Committee London - City & East; 13/LO/1948), is a set of three-dimensional SW images of newborns affected by the hypoxic-ischemic encephalopathy (HIE). This dataset

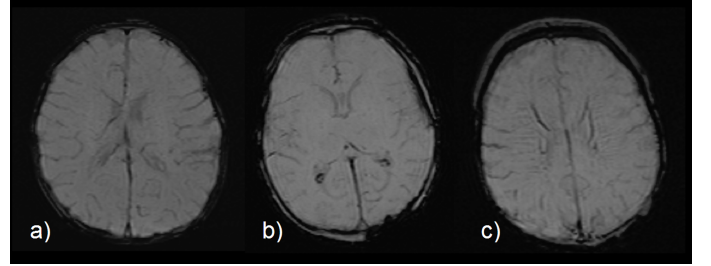


Figure 8: Susceptibility-weighted example images from our clinical dataset (NRES Committee London - City & East; 13/LO/1948): a) patient not affected by HIE (under oxygen during MRI acquisition), b) patient affected by the HIE (under oxygen), c) patient affected by the HIE (not under oxygen)

is composed of 17 patients where 7 are affected by HIE, and 10 for which there was no clinical evidence for HIE. The oxygenated samples are 10 out of 17 where the presence of HIE is evenly distributed in both classes. We assume that the textures in our medical data to be approximately isotropic. The size of the images for this dataset are within  $256H \times 176W \times 48D$  and  $256H \times 256W \times 64D$  voxels. Figure 8 shows some example images from our set of three-dimensional scans.

We convert all the images in the dataset using different extended 3D LBP and we extract a histogram of LBP codes for each patient. These histograms are then introduced in a simple learning machine (classifier) to emphasize the discriminatory power of our descriptors. One can argue that a simple comparison of histograms derived from the grey-scale voxels are enough to distinguish the two categories. In fact, this method has a near zero discriminatory power because it discards the



Table 8: Classification accuracies [%] discriminating between oxygenated (patient that were under oxygen during MRI acquisition) and non-oxygenated samples from our clinical dataset (NRES Committee London - City & East; 13/LO/1948).

$\{\downarrow\} - LBP_{\{\rightarrow\}}^{riu3g}$	Spatial resolution		
	(1,42)	(2,92)	(3,252)
<i>NI</i>	<b>89.9</b>	65.9	49.5
<i>RD</i>	82.5	81.1	87.5
<i>NI/RD/CI</i>	88.2	61.5	60.6
<i>NI + RD + CI</i>	82.2	66.6	65.3

interaction between voxels and it is highly sensitive to uneven intensity variations.

A K-nearest-neighbours classifier (L1 metric) is employed here with  $k = 3$  in order to reduce a possible overfitting due to the small number of samples in our dataset. The descriptors employed here are the *NI - LBP* at different resolutions, the *RD - LBP*, the joint *NI/RD/CI - LBP* and the concatenated *NI + RD + CI - LBP*.

The classification accuracies presented in Table 8 demonstrate that a difference in the texture between oxygenated and non-oxygenated MR images exist, especially by employing our *NI - LBP<sup>riu3g</sup>* at the spatial resolution of (1,42). This descriptor can extract salient information from the interaction of the pixels/voxels in a more accurate way than what human perception can do.

#### 4. Conclusion

This paper proposes a new set of three-dimensional fully rotation invariant texture descriptors exploiting existing features developed for two-dimensional LBP (Liu et al., 2012) and a three-dimensional rotation invariant LBP (Paulhac et al., 2008). Among the descriptors proposed here, the *NI - LBP<sup>riu3g</sup>* demonstrate an excellent discriminatory power on a dataset of synthetic textures followed by the *RD - LBP<sup>riu3g</sup>*. In the experiments we confirm the utility of adding a third dimension to local binary patterns as demonstrated by Paulhac et al. (2008). We show also that a combination of local binary patterns in various scales leads in most cases to better results. Furthermore, in our experiments we demonstrate that joint combinations of histograms and simple concatenations lead to nearly similar results, however, in the second case, the shorter histograms are more efficient in applications where fast computations are required.

#### References

Ahonen, T., Hadid, A., Pietikäinen, M., 2006. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 28, 2037–2041.

Banerjee, J., Moelker, A., Niessen, W., van Walsum, T., 2013. 3d lbp-based rotationally invariant region description. In: Park, J., Kim, J. (Eds.), *Computer Vis.-ACCV 2012 Workshops*. Springer, Daejeon, Korea, pp. 26–37.

Bayramoglu, N., Zhao, G., Pietikäinen, M., 2013. Cs-3dlbp and geometry based person independent 3d facial action unit detection. In: *International Conference on Biometrics (ICB)*, 2013 June 4. IEEE, Madrid, Spain, pp. 1–6.

Biasotti, S., Cerri, A., Aono, M., Ben Hamza, A., Garro, V., Giachetti, A., Giorgi, D., Godil, A., Li, C., Sanada, C., Spagnuolo, M., Tatsuma, A., Velasco-Forero, S., 2016. Retrieval and classification methods for textured 3d models: a comparative study. *The Visual Computer* 32, 217–241.

Depeursinge, A., Foncubierta-Rodriguez, A., Van De Ville, D., Müller, H., 2014. Three-dimensional solid texture analysis in biomedical imaging: Review and opportunities. *Méd. Image Analysis* 18, 176–196.

Fehr, J., Burkhardt, H., 2008. 3d rotation invariant local binary patterns. In: *The 19th International Conference on Pattern Recognit., ICPR 2008*. IEEE, Tampa, FL, pp. 1–4.

Harwood, D., Ojala, T., Pietikäinen, M., Kelman, S., Davis, L., 1995. Texture classification by center-symmetric auto-correlation, using kullback discrimination of distributions. *Pattern Recognit. Letters* 16, 1–10.

Huang, Y., Wang, Y., Tan, T., 2006. Combining statistics of geometrical and correlative features for 3d face recognition. In: *British Machine Vision Conference*. Edinburgh, UK, pp. 879–888.

Huynh, T., Min, R., Dugelay, J., 2012. An efficient lbp-based descriptor for facial depth images applied to gender recognition using rgb-d face data. In: *ACCV 2012, Workshop on Computer Vision with Local Binary Pattern Variants*. Daejeon, Korea.

Kenner, H., 2003. *Geodesic math and how to use it*. University of California Press, Berkeley.

Kovalev, V., Kruggel, F., Gertz, H., von Cramon, D., 2001. Three-dimensional texture analysis of mri brain datasets. *IEEE Transactions on Méd. Imaging* 20, 424–433.

Li, S., Zhao, C., Ao, M., Lei, Z., 2005. Learning to fuse 3d+2d based face recognition at both feature and decision levels. *Analysis and Modelling of Faces and Gestures, Lecture Notes in Computer Science* 3723, 44–54.

Liu, L., Zhao, L., Long, Y., Kuang, G., Fieguth, P., 2012. Extended local binary patterns for texture classification. *Image and Vis. Computing* 30, 86–99.

Liu, S., Cai, W., Wen, L., Feng, D., 2011. Volumetric congruent local binary patterns for 3d neurological image retrieval. In: *26th International Conference Image and Vision Computing New Zealand (IVCNZ 2011)*. pp. 272–276.

Majtner, T., Svoboda, D., 2014. Comparison of 3d texture-based image descriptors in fluorescence microscopy. In: *16th International Workshop, IW-CIA 2014, May 28-30, 2014*. Proceedings. Springer International Publishing, Brno, Czech Republic, pp. 186–195.

Ojala, T., Pietikäinen, M., Harwood, D., 1996. A comparative study of texture measures with classification based on feature distributions. *Pattern Recognit.* 29, 51–59.

Ojala, T., Pietikäinen, M., Mäenpää, T., 2002. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24, 971–987.

Paulhac, L., Makris, P., Ramel, J., 2008. Comparison between 2d and 3d local binary pattern methods for characterisation of three-dimensional textures. In: *Campilho, A., Kamel, M. (Eds.), Image analysis and recognit., ICIAR 2008*. Springer, Póvoa de Varzim, Portugal, pp. 670–679.

Paulhac, L., Makris, P., Ramel, J., 2009. A solid texture database for segmentation and classification experiments. In: *VISAPP 2009 - Proc. of the Fourth International Conference on Computer Vis. Theory and Applications Volume 2*. INSTICC Press, Lisboa, Portugal, pp. 135–141.

Sorensen, L., Shaker, S., de Bruijne, M., 2010. Quantitative analysis of pulmonary emphysema using local binary patterns. *IEEE Transactions on Méd. Imaging* 29, 559–569.

Werghi, N., Berretti, S., Del Bimbo, A., 2015a. The mesh-lbp: A framework for extracting local binary patterns from discrete manifolds. *IEEE Transactions on Image Processing* 24, 220–235.

Werghi, N., Berretti, S., Del Bimbo, A., 2015b. Representing 3d texture on mesh manifolds for retrieval and recognition applications. In: *Computer Vision and Pattern Recognition (CVPR)*, 2015 IEEE Conference on. IEEE, Boston, Massachusetts, pp. 2521–2530.

Werghi, N., Berretti, S., Del Bimbo, A., Pala, P., 2013. The mesh-lbp: Computing local binary patterns on discrete manifolds. In: *Computer Vision Workshops (ICCVW)*, 2013 IEEE International Conference on. IEEE, Sydney, Australia, pp. 562–569.

Zhao, G., Pietikäinen, M., 2006. Local binary pattern descriptors for dynamic texture recognition. In: *Pattern Recognition*, 2006. ICPR 2006. 18th Inter-

- national Conference on. Vol. 2. IEEE, Hong Kong, China, pp. 211–214.
- Zhao, G., Pietikäinen, M., 2007a. Dynamic texture recognition using local binary patterns with an application to facial expressions. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29, 915–928.
- Zhao, G., Pietikäinen, M., 2007b. Dynamic texture recognition using volume local binary patterns. In: *WDV 2005 and WDV 2006 Workshop on Dynamical Vision*. Vol. 4358. Springer, Beijing, China, pp. 165–177.

# Extended three-dimensional rotation invariant Local Binary Patterns

## Image and Vision Computing IMAVIS-D-16-00270

### Response to Reviewers

Leonardo CITRARO, ldo.citraro@gmail.com

February 28, 2017

## 1 Response to the valuable comments of the reviewers

---

**Reviewer #2:** Authors have addressed most of the previous comments that was made for the first review. However, original theoretical contribution is still limited and there are still missing points which needs to be done before publishing this paper.

Before, I suggested to make a comparison with the method proposed by Banerjee et al. (2012) with increased descriptor radius and sampling rate. The authors experimented the method only with increased sampling rate. Descriptor radius should also be increased. Instead, they have commented the following: "We understand that some descriptors may perform better if we carefully select the right radius and sampling scheme but testing all possible combinations (considering also many other possible combinations for our own descriptors) is really time consuming and out of the scope for this paper." Certainly I did not mean to select all possible combinations of radius and sampling rate but at least 2-3 different radii must have been experimented. The aim is to make a fair comparison. The proposed method also depends on several parameters and reader needs to know which one performs better at different conditions.

Response: Corrected. In order to make a fair comparison of the methods proposed by Banerjee et al. (2012), Fehr and Burkhardt (2008) and our methods, we performed experiments with different spatial resolutions for all of them. Moreover, for two methods exploiting the spherical harmonics, we decided to analyse the effect of the number of bins in the histograms on the scores. Our previous assumption of using a carefully chosen number of bins for the resulting histograms was erroneous. Tables 2 and 3 show the results for the Banerjee et al. (2012) and Fehr and Burkhardt (2008) methods with different number of bins in the histogram and different spatial resolutions. Figure 6 shows the most important scores of all descriptors tested. Due to the numerous changes made, Section 3.5 (Experiment 1) has been now re-factored and rewritten almost entirely. Most of the previous reasoning are still present but in different order.

During these experiments we identified an error in the code concerning the geometry of our 3D  $RD - LBP_{3,252}$ . Results related with this descriptor are now slightly better; Table 5, 6, 7 and 8 are affected as well as Figure 7 have slightly changed but the reasoning made in the discussion and conclusion in the previous paper draft still hold.

Regarding the theory, we have now added a paragraph (third paragraph) in section 2.2 explaining why the NI-LBP was created and why the combination of these extended features outperformed the Ojala pair LBP/VAR.

Naming conventions for different methods are still difficult to read especially when examining the experimental results from tables and charts. Each time the reader needs to check the text to find about the method. References (and "proposed") should also be written alongside.

Response: Corrected. We agree with the fact that some name conventions are long and difficult at first sight but at the same time it is difficult to simplify them without losing important information. In order to improve the readability of the paper we shortened as possible some method's names and we summarised them in Table 1 (Table of abbreviations). Moreover, in the text we better referenced the descriptors with authors names and where possible we added keywords to identify the methods.

Authors have commented on the reasons why they do not make a comparison with 3D convolutional neural networks. I could not accept all of the excuses but I recommend at least to write a less strong conclusion. The following statement becomes unreliable if the CNNs are not considered as state of the art descriptors: "Our results demonstrate the effectiveness of

the extended LBPs and improvements against other state-of-the-art three-dimensional texture descriptors on this dataset.”

Response: Corrected. We modified the Abstract by saying that our paper compares hand-crafted features. In the abstract: ”Our results demonstrate the effectiveness of the extended LBPs and improvements against other state-of-the-art hand-crafted three-dimensional texture descriptors on this dataset.”

Also in Section 3.5 (Experiment 1) we softened our statement (penultimate paragraph on page 8): ”... standing among the top state-of-the-art hand-crafted three-dimensional texture descriptors.”

## 2 Modifications

---

- **Abstract:** we softened the conclusion as follow: ”Our results demonstrate the effectiveness of the extended LBPs and improvements against other state-of-the-art hand-crafted three-dimensional texture descriptors on this dataset.”
- **3.3 Variable selection:** The description of the experiments for the two methods employing the spherical harmonics have slightly changed in order to explain that we tested multiple histogram length and spatial resolutions.
- **3.5 Experiment 1:** we re-factored and rewritten most of this section due to the numerous additional experiments made.