A SCALE-ADAPTIVE EXTENSION TO METHODS BASED ON LBP USING SCALE-NORMALIZED LAPLACIAN OF GAUSSIAN EXTREMA IN SCALE-SPACE

Sebastian Hegenbart and Andreas Uhl

University of Salzburg
Department of Computer Sciences
Salzburg, Austria

ABSTRACT

Local Binary Patterns and its derivatives have been widely used in the field of texture recognition over the last decade. A restriction of methods based on LBP is the variance in terms of signal scaling. This is mainly caused by the fixed LBP radius and the fixed support area of sampling points. In this work we present a general framework to enhance the scale-invariance of all LBP flavored methods, which can be applied to existing methods with minimal effort. Based on scale-normalized Laplacian of Gaussian extrema in scale-space, the global scale of a texture in question is estimated, combined with a confidence measure, to compute scale adapted patterns. By using the notion of intrinsic scales, textures are analyzed at appropriate LBP scales. A comprehensive experimental study shows that the scale-invariance of three different LBP based methods (LBP, LTP, Fuzzy LBP) is highly improved by the proposed extension.

Index Terms— scale, adaptive, LBP, scale-space, estimation

1. INTRODUCTION

In certain scenarios, medical imaging for example, textures are captured at various perspectives and distances [1]. These variations caused by camera motion lead to a visualization of textures under different scales. Methods that are invariant in terms of signal scale can therefore improve the accuracy of an automated classification in such a setting.

Since the introduction of the Local Binary Patterns (LBP) method [2], a variety of LBP based flavors have been developed and applied in various specialized texture recognition scenarios. All LBP based methods share the limitation of being highly affected by scaling of a signal however. Ojala and Mäenpää introduced multi-resolution Local Binary Patterns [3], using a set of different radii with appropriate sampling areas. While this approach improves the discriminative power of the method, it does not employ a scale selection mechanism and hence does not improve invariance in terms of signal scaling.

The idea of combining scale-space extrema with LBP to improve scale-invariance has also been explored by Li et al. [4]. Their approach utilizes the scale of a scale-space maxima at a pixel position as the scale of the Local Binary Pattern descriptor, using a fixed number of neighbors (8) with a fixed sized neighbor sampling area.

Our experimentation has shown that scale selection, based on a single pixel location, is very prone to error, especially for non-regular textures such as shown in Figure 1. We therefore compute a global scale estimation combined with a confidence measure for the estimation to compute scale adapted patterns along a fixed grid (all pixel positions) in an image. Experimental data also shows that a direct mapping from a scale in scale-space to LBP scale is far from optimal. The fact that the estimated scale at a pixel level highly correlates to the intrinsic scale of a texture, leads to rather large LBP scales if using a direct mapping. This reduces the discriminative power of the LBP patterns due to the decreased correlation between sampling points and reference point and leads to sparse sampling if fixed sampling area dimensions are used, reducing the discriminative power even further. We solve this problem by introducing the notion of intrinsic scale, computing a mapping from estimated scale in scale-space to a much more suitable LBP scale. We adjust the sampling area dimensions according to the adapted LBP scale to improve the scale-invariance even further. In this work a general framework to enhance the scale-invariance of all LBP flavored methods, which can be applied to existing methods with minimal effort, is presented.

2. A SCALE-ADAPTIVE EXTENSION TO LBP BASED METHODS

The scale-space theory was first extensively explored in the field of signal processing by Lindeberg [5, 6]. It presents a framework to analyze signals at different scales. Let $f : \mathbb{R}^2 \mapsto \mathbb{R}$ represent a continuous signal, then the linear scale-space representation $L : \mathbb{R}^2 \times \mathbb{R}_+ \mapsto \mathbb{R}$ is defined by

$$L(\cdot; \sigma) = g(\cdot; \sigma) * f, \qquad (1)$$
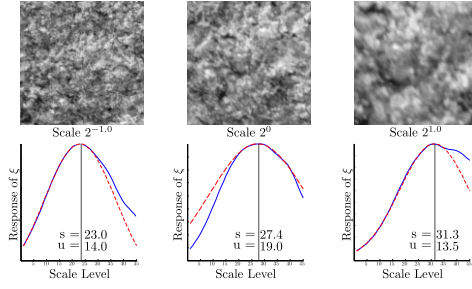
Fig. 1: Scale Estimation of a non-Regular Texture (stone2)

with initial condition $L(\cdot; 0) = f$. Where $\sigma \in \mathbb{R}_+$ is the scale parameter, g is a Gaussian function and "$*$" denotes convolution. The scale-space family L is the solution to the diffusion equation (heat equation):

$$\partial_\sigma L = \sigma \left(\frac{\partial^2 L}{\partial x^2} + \frac{\partial^2 L}{\partial y^2} \right) = \sigma \triangle L. \quad (2)$$

We construct the scale-space and compute the scale-normalized Laplacians ($\sigma^2 |\triangle L(\cdot; \sigma)|$, denoted as $\overline{\triangle} I(\cdot; \sigma)$) of each image I at each location $x \in \mathbb{N}^2$ at different scales with $\sigma = c\sqrt{2}^k, k \in \{-4, -3.75, \ldots, 7.75, 8\}$ and $c = 2.1214$. Note that the parameter c acts as a scaling factor of the scale-space and was initially chosen such that the center scale equals a 3 pixel radius. We however found during experimentation that the intrinsic scale of natural textures tends to be large. We therefore expanded the scale-space to cover larger scales as well.

Methods based on scale selection employing the scale-space abstraction identify image locations which are simultaneously a local extremum with respect to both the spatial coordinates and the scale-space parameter. Hegenbart et al. [1] use a local scale reliability mask to improve the reliability of the scale estimation based on such extrema. Experimentation showed however that the utilization of such locations to compute LBP based feature vectors in general leads to an insufficient number of computed patterns and a reduced discriminative power of the feature. The scale selection based on a single pixel location, such as performed by Li et al. [4], however is prone to error, as not each pixel is at a representative scale, especially for non-regular textures as shown in Figure 1. As a consequence we compute a global scale estimate combined with an uncertainty for the estimation to compute LBP patterns along a fixed grid (all pixel positions) in an image, adapted to the global scale of the texture. Let δ denote the Kronecker delta, the scale estimation response function ξ is then

$$\xi(t) := \sum_{x,y} \delta(\arg \max_{\sigma}(\overline{\triangle} L(x, y; \sigma)), t) \overline{\triangle} L(x, y; t). \quad (3)$$

The global scale is identified by searching for the first local maximum of ξ which is then used as seed point for a least-squares Gaussian fit. By using the first local maximum we are capable of consistently estimating the scale of textures exhibiting more than a single dominant scale. The quality of the estimation is improved by considering only data points within a certain offset from the seed point. In our implementation an offset of ± 5 scales is used to fit the Gaussian function. Finally the average value (s) of the fitted Gaussian function is interpreted as the estimated scale where the standard deviation is used as uncertainty of the estimation (u). Due to the fact that the accuracy of the scale estimation is not perfect, we extract weighted LBP patterns at multiple scales to improve robustness, taking the uncertainty of the estimation into account. The weighted patterns contribute to the LBP histogram, based on the response of the unnormalized Gaussian function at the specific scale level. In our implementation only scale levels with a response ≥ 0.9 were used. Figure 1 illustrates the fitted Gaussian function (dashed line) to the scale estimation response function (solid line) of three textures at different scales.

2.1. Intrinsic Texture Scale

The response of the scale-normalized Laplacian of Gaussians (LoG) attains a maximum if the zeros are aligned with a circular shaped structure. Hence scales estimated, based on the LoG, correlate strongly with the scale of the dominant circular shaped structures of a texture. As a consequence, the estimated scale is highly related to an essential property of each texture, the *intrinsic scale* of a texture.

A texture exhibiting pebbles for example and a texture exhibiting sand, captured at the same distance, might have equal scales in terms of camera-scale, but different scales in terms of the scale-space, a consequence of different intrinsic scales. In contrast, sand and pebbles captured at a different camera-scales, corresponding to the difference of the textures' intrinsic scales, are equal in scale in terms of the scale-space. Scales estimated in the scale-space domain are therefore always a combination of the intrinsic texture scale and the camera-scale.

Utilizing LBP based methods, textures are described by the means of the joint distribution of underlying micro structures. The discriminative power is not directly related to the scale of the dominant structures of an image. This statement is based on the observation that LBP based methods are successfully used in classification scenarios with multiple textures and multiple different intrinsic scales, using a set of fixed sized LBP radii. Hence, a direct mapping between estimated scale in scale-space of a texture to LBP scale, introduces several problems as discussed in Section 1, without improving the descriptive capabilities of the method in general.

The identification of an intrinsic scale of a general texture is a non-trivial problem. A requirement on an intrinsic scale estimation method would be scale-invariance, a property that the LoG response in scale-space does not provide.

Based on the property that the intrinsic scale of a texture is scale-invariant, the intrinsic factors cancel each other out for two estimated scales in scale-space of the same texture. We hence estimate the scale in scale-space per texture class in the training data, denoted as *trained base scale*, as the median of all estimated scales from all images within a class.

This approach requires that all samples of a specific texture class are at a single or relatively close camera-scale in the training data. This requirement could be loosened by identifying the *trained base scale* per class and camera-scale however. A benefit of estimating the *trained base scale* per class is, that additional information such as a shape model per texture class could be computed and used for improving the feature extraction further. This is part of our current work.

By using the *trained base scale* of a texture class, we can define a mapping from the scale-space domain to the LBP scale domain (the LBP radius). For a texture with estimated scale in scale-space s the adapted LBP radius is then computed in reference to the *trained base scale* ($\bar{s}_l$) of texture class l as

$$\iota(s,l) = b\frac{s}{\bar{s}_l}, \tag{4}$$

with b denoting the defined LBP radius at the *trained base scale* $\bar{s}_l$. Please note that the linearity of this mapping is a requirement for scale-invariance. The value of b defines the LBP scale the training textures are analyzed at. In order to be able to adapt to down-scaled textures, the value requires to be larger than the minimal LBP radius. We use $b = 3$ as default.

Considering the extraction of feature vectors for evaluation, we are not capable of identifying the corresponding *trained base scale* for such input textures, due to the inability of estimating the intrinsic scale of the texture. Hence, for each input texture a set of feature vectors is computed, one feature vector in relation to the specific *trained base scale* of each class in the training set. Feature vectors computed in relation to the same texture class will be based on a matching *trained base scales* (the input sample and the texture class are at the same intrinsic scale), canceling out the intrinsic scale factors. Feature vectors computed in relation to other texture classes (and other *trained base scales*) are computed at the wrong relative LBP scale.

The pairwise comparison of feature vectors computed at different *trained base scales* can lead to very high LBP scales. As a consequence such feature vectors exhibit a higher general similarity among all textures (due to the high amount of low-pass filtering), leading to a decreased discriminative power of the system. As a consequence, only features computed at the same *trained base scale* are compared during classification. Please note, that this poses no unfair bias or advantage to the classifier as each of the feature vectors of an evaluation sample is compared to the corresponding feature vector of each class in the training data.
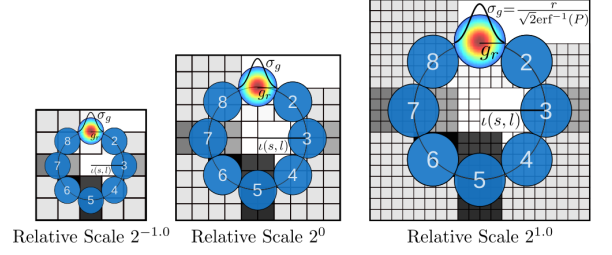

Fig. 2: Adaptive Radius and Adaptive Sampling Areas

2.2. Adaptive Sampling Support Area Dimension

Scaling of a texture leads to a scaled spatial extent of structures. Therefore the number of pixels covering structural information changes. As a consequence, the size of the sampling support area has to be adapted accordingly. Ojala and Mäenpää [3] first used Gaussian filtering to adapt the sampling support area to various LBP scales to compute multi-resolution LBP, generally improving the robustness of the method. By employing low-pass filtering, a pixel at a single spatial location encodes information of its spatial neighborhood. We use the estimated global scale of a texture in relation to a *trained base scale* to adapt the LBP radius as well as the size of the sampling support area to achieve scale-invariance. The radius of the Gaussian filter for a texture at estimated global scale s in relation to the *trained base scale* of texture l is computed as

$$g_r = \frac{\iota(s,l)\pi}{N}, \tag{5}$$

for N being the number of neighbors. The Gaussian filter coefficients are then computed such that P percent of the mass of the Gaussian function is covered

$$\int_{-g_r}^{g_r} e^{-\frac{x^2}{2\sigma_g^2}} dx = P \int_{-\infty}^{\infty} e^{-\frac{x^2}{2\sigma_g^2}} dx$$
$$2\int_{0}^{g_r} e^{-\frac{x^2}{2\sigma_g^2}} dx = P\sigma_g\sqrt{2\pi}$$
$$\sigma_g = \frac{g_r}{\sqrt{2}\mathrm{erf}^{-1}(P)}. \tag{6}$$

We chose P to be 0.99 which corresponds to 99% of the mass of the Gaussian function. As the sampling of a Gaussian function with very few sampling points leads to a large error we use the error function (erf) to improve the stability of the computation of the one dimensional Gaussian filters centered at 0

$$G(x;\sigma_g) = \frac{-\mathrm{erf}(\frac{x-0.5}{\sigma_g}) - \mathrm{erf}(\frac{x+0.5}{\sigma_g})}{2}, \tag{7}$$

which are then used in a separable convolution. Note that, as a bonus, the computed filter is already normalized, therefore the re-normalization can be avoided. Figure 2 illustrates the relation between estimated scale in scale-space and the adapted LBP scale.
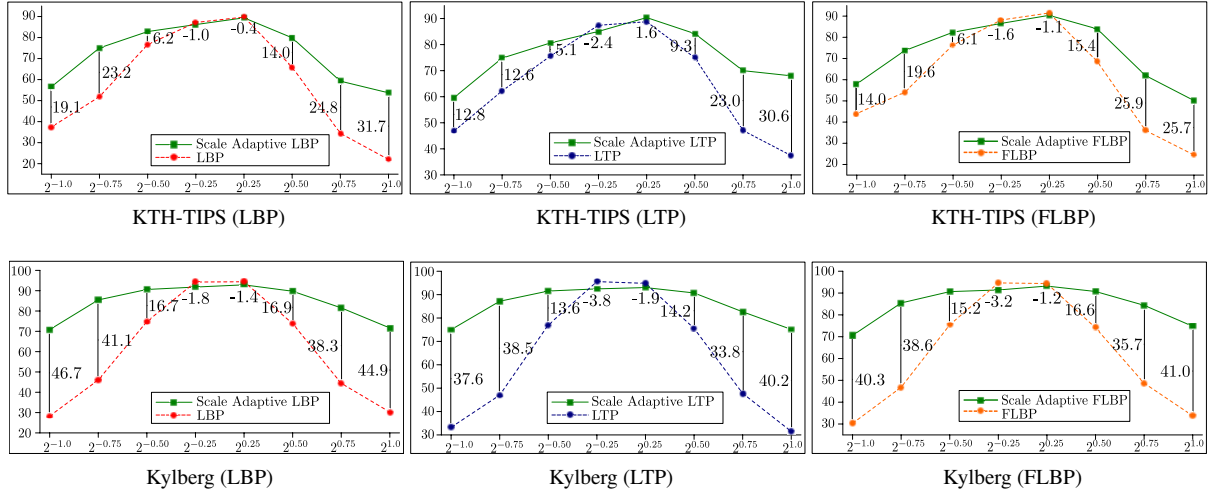
Fig. 3: Mean Overall Classification Accuracies

3. EXPERIMENTS AND RESULTS

We constructed a set of experiments to analyze the impact of the proposed scale-adaptive extension to three different, popular types of LBP based methods, which all utilize different types of encoding schemes. The standard LBP method [2], the Local Ternary Patterns (LTP) operator [7] as well as the Fuzzy Local Binary Patterns (FLBP) method [8] are compared with the scale-adapted variation of each of those methods. Please note that the standard methods were used in combination with the multi-resolution Local Binary Patterns extension [3] using three scales and 8 neighbors, the best configuration we were able to find for the given datasets.

The experiments are based on two different texture databases. The KTH-TIPS database [9] exhibits texture images from 10 different materials captured at 9 different scales with 9 samples per material. Sub-images of size 128×128 pixels were extracted from the center of each original image. Unfortunately, besides KTH-TIPS there are no other publicly available high quality texture databases with an available ground-truth of scales. We therefore had to resort to a simulation of the scaling of textures. A subset of the Kylberg texture database [10], consisting of 28 materials with 160 unique texture patches per class, captured at a single scale, was used for the simulation. The high resolution of each patch (576×576 pixels) allowed us to simulate the scaling without relying on up-sampling, leading to a smaller amount of interpolation artifacts. The simulation of scaling was performed according to the scales of the KTH-TIPS database, interpreting the original image patches as the maximum scale $2^{1.0}$. Image patches of size 128×128 pixels were then extracted from the center of the re-scaled patches. Due to the huge number of samples in the Kylberg database we use a subset consisting of 20 unique texture patches per class (5 per image) for experimentation.

The experiments were designed to explicitly reflect the scale invariance properties of the studied methods. We chose KTH-TIPS scale 5 and Kylberg scale 2^0 as the training scale. This gives us the opportunity to study the method's capability of adapting to higher as well as lower relative scales.

The classification was performed based on a k-nearest neighbors classifier, utilizing the histogram intersection as similarity metric. The minimum number of neighbors was set to 1 for all experiments while the maximum number of neighbors was set according to the maximum number of samples per texture class (9 for KTH-TIPS and 20 for Kylberg).

Figure 3 shows the mean overall classification accuracy (OCR) over all k-values. The numbers between the results illustrate the absolute difference in mean OCR between the scale-adapted type of a method and the standard version. The horizontal axes denote the relative scale differences as compared to the training set.

4. DISCUSSION AND CONCLUSION

The experimental results show that the scale-adaptive extension improved the scale-invariance of all three flavors of the LBP method. The scale-adapted methods performed slightly worse as compared their standard counterpart at very small relative scale differences. This is caused by some failed scale estimations. Considering large scale differences however, the scale-adapted methods vastly outperform the standard methods, with improvements of over 40 percentage points. Although the scale-adaptive extension comes at the cost of higher computational demand, it can improve the classification accuracy of LBP based methods in a setting with varying scales significantly without changing the encoding or extraction scheme of the standard method, and could therefore be used in combination with a variety of LBP based methods.

5. REFERENCES

[1] S. Hegenbart, A. Uhl, A. Vécsei, and G. Wimmer, "Scale invariant texture descriptors for classifying celiac disease," *Medical Image Analysis*, vol. 17, no. 4, pp. 458 – 474, 2013.

[2] T. Ojala, M. Pietikäinen, and D. Harwood, "A comparative study of texture measures with classification based on feature distributions," *Pattern Recognition*, vol. 29, no. 1, pp. 51–59, January 1996.

[3] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971–987, July 2002.

[4] Z. Li, G. Liu, Y. Yang, and J. You, "Scale- and rotation-invariant local binary pattern using scale-adaptive texton and subuniform-based circular shift," *IEEE Transactions on Image Processing*, vol. 21, no. 4, pp. 2130–2140, 2012.

[5] T. Lindeberg, "Scale-space for discrete signals," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12, no. 3, pp. 234–254, 1990.

[6] T. Lindeberg, *Discrete scale-space theory and the scale-space primal sketch*, Ph.D. thesis, Royal Institute of Technology, 1991.

[7] X. Tan and B. Triggs, "Enhanced local texture feature sets for face recognition under difficult lighting conditions," in *Analysis and Modelling of Faces and Gestures*, October 2007, vol. 4778 of *Lecture Notes in Computer Science*, pp. 168–182.

[8] D. Iakovidis, E. Keramidas, and D. Maroulis, "Fuzzy local binary patterns for ultrasound texture characterization," in *ICIAR*. 2008, vol. 5112 of *Lecture Notes in Computer Science*, pp. 750–759, Springer.

[9] E. Hayman, B. Caputo, M. Fritz, and J.-O. Eklundh, "On the significance of real-world conditions for material classification," in *Proceedings of the 8th European Conference on Computer Vision (ECCV)*. 2004, vol. 3024 of *Lecture Notes in Computer Science*, pp. 253–266, Springer.

[10] G. Kylberg, "The kylberg texture dataset v. 1.0," External report (Blue series) 35, Center for Image Analysis, Swedish University of Agricultural Sciences, Uppsala University, Uppsala, Sweden, September 2011.