Scale-Adaptive Texture Classification

Michael Gadermayr
University of Salzburg
Department of Computer Sciences
Email: mgadermayr@cosy.sbg.ac.at

Sebastian Hegenbart
University of Salzburg
Department of Computer Sciences
Email: shegen@cosy.sbg.ac.at

Andreas Uhl
University of Salzburg
Department of Computer Sciences
Email: uhl@cosy.sbg.ac.at

Abstract—Scale invariant texture analysis is a fundamental challenge in image processing. As a consequence of the scale invariance, these kind of features are often characterized by a lower discriminative power. We observed, that scale invariant features did not pose a benefit in classification scenarios with varying scales in the training set. This is supposed to be an effect caused by an implicit scale selection done by the classification method. In this work, we analyze this effect based on the k-nearest neighbor classifier. Inspired by this effect, we employ global scale estimation algorithm utilizing scale-normalized Laplacian of Gaussian extrema in scale space, to improve the classification accuracies of scale variant features in a scenario with varying scales. We propose a general framework for scale-adaptive classification, which proved to improve the classification accuracies with a variety of feature extraction methods in such a scenario.

I. INTRODUCTION

Scale invariant texture analysis is a fundamental challenge in image processing. This theoretically alleviates the restriction that texture data has to be acquired at a constant distance to the camera, a property which is highly relevant in practice. In order to achieve scale invariance, various approaches exist. In Zhang et al. [1], a review about invariant texture analysis methods is presented. Another survey of Hegenbart et al. [2] investigates scale invariance in combination with medical image data. However, experiments have shown that many features that are not explicitly scale invariant [3], [4], [5], [6], [7], [8], [9], are performing effectively in terms of classification accuracy in scenarios with variably scaled textures. We observed, that scale invariant features in general do not outperform these scale variant features, a behavior caused by a decrease of the discriminative power due to the invariance. The rationale for using scale-invariant feature is based on the assumption that the intra-class variability within a scale of a texture generally is lower than the intra-class variability over multiple scales. As a consequence, scale-variant features exhibit a higher intra-class variability over multiple scales as compared to scale-invariant features. Therefore, a pair of feature vectors with the same scale of a texture class has a higher probability to be closer (in terms of the feature distance) as compared to feature vectors of a texture class at different scales. This property implicitly leads to a kind of scale selection during the classification process when using scale-variant features. Especially, the partitioning performed by the k-nearest neighbor classifier increases this effect.

In this work, we particularly investigate the implicit scale selection mechanism of the k-nearest neighbor method. In a large experimental setup we evaluate this effect using a variety of scale-variant and scale-invariant feature extraction methods. Moreover, we propose a framework to increase this beneficial effect. Based on a global scale estimation algorithm utilizing scale-normalized Laplacian of Gaussian extrema in scale space we utilize the superior discriminative power of scale-variant features in a scale-adaptive classification.

This paper is structured as follows. We describe the scale estimation algorithm in Section II. The proposed scale-adaptive classification framework is covered in Section III. The conducted experiments are explained and discussed in Section IV. Section V finally concludes this paper.

II. SCALE ESTIMATION

We employ a global scale estimation algorithm [10] which is based on scale-normalized Laplacian of Gaussian extrema in scale-space. The scale-space theory was first extensively explored in the field of signal processing by Lindeberg [11].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

$$L(\cdot; \sigma) = g(\cdot; \sigma) * f, \qquad (1)$$

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):

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

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 (3D maxima), a prominent example is the Scale Invariant Feature Transform (SIFT [12]).

Experimentation has shown however that the utilization of such locations for a global scale estimation is unreliable. This can be seen in Figure 1, comparing the distribution of
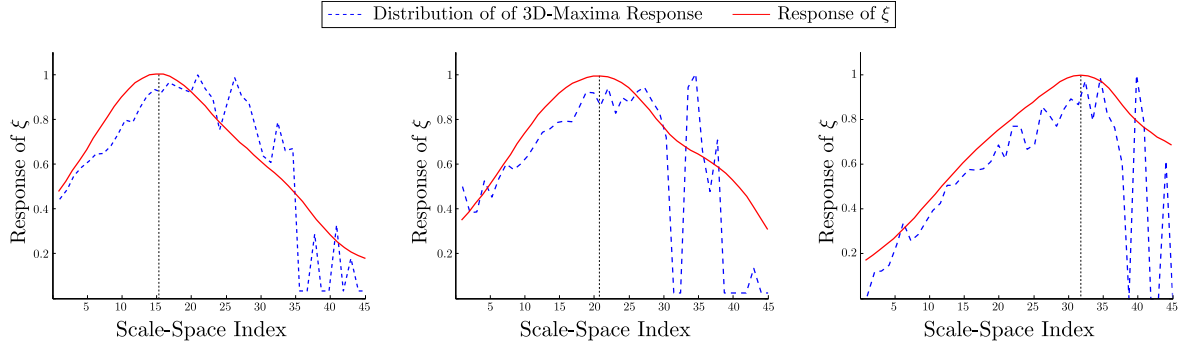
Fig. 1: Distribution of 3D-Maxima compared to the Response of ξ.

the responses of the 3D maxima with the responses of the scale estimation function ξ for three scaled versions of an image of the KTH-TIPS database. The x-axis denotes the scale level within scale space and the y-axis denotes the average normalized response at the specific scale level. This property especially holds in the case of non-regular textures as shown in Figure 2. Considering this type of textures, the extrema are either at various different scales or only a small number of extrema is present, leading to unreliable scale estimations. We therefore use the distribution of responses of the scale normalized Laplacians to estimate a global scale. The scale estimation response function ξ is

$$\xi(t) := \sum_{x,y} \overline{\triangle} I(x,y;t). \tag{3}$$

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 scale levels is used to fit the Gaussian function. Finally the mean value $\tilde{s}$ of the fitted Gaussian function is interpreted as the dominant level in scale-space. The standard

deviation u of the fitted Gaussian is used as uncertainty of the estimation. Figures 2 and 3 illustrate the fitted Gaussian function (dashed line) to the scale estimation response function (solid line) of three textures at different scales.

The response of the scale-normalized Laplacian of Gaussian (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.

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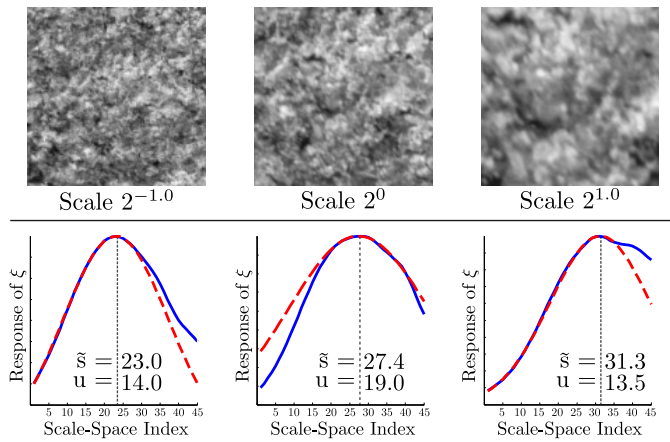

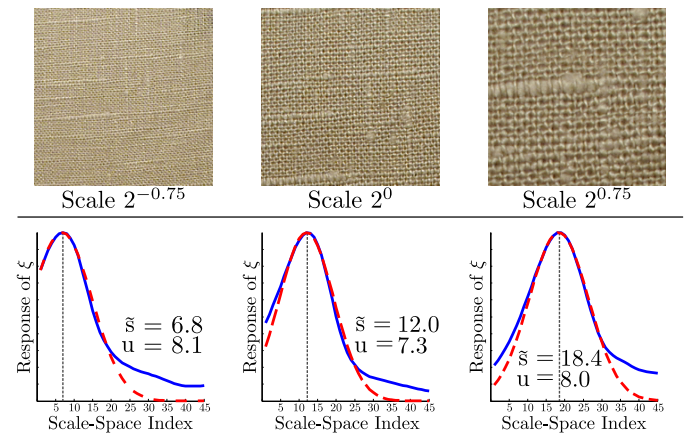Fig. 2: Scale Estimation of a non-Regular Texture (stone2).

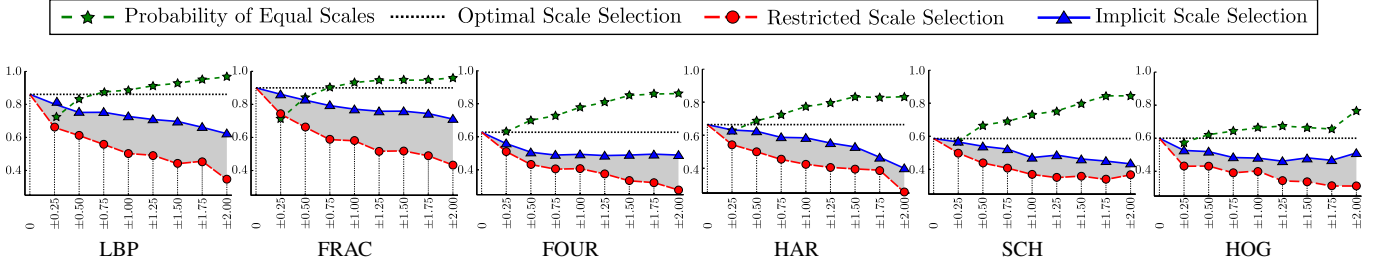Fig. 3: Scale Estimation of a non-Regular Texture (linen).

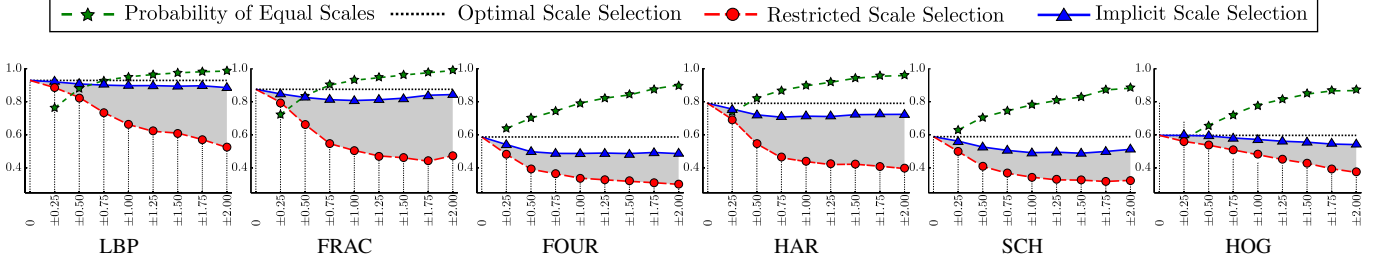Fig. 4: Analysis of Implicit Scale Selection using the KTH-TIPS Database.

Fig. 5: Analysis of Implicit Scale Selection using the Kylberg Database.

of camera-scale and intrinsic texture scale.

A. Performance

We evaluated the performance of the scale estimation algorithm to assess the practicality of the method in an experimental setup with a huge number of data. We estimated the scales based on a data set comprised of 40300 images with a dimension of 128×128 pixels each.

Our implementation was written in Java (JDK 8) utilizing parallel processing. The experiment was conducted on an i5-2400 CPU with 4 cores running at 3.10 GHZ. We assume that an implementation written in a language that is compiled to binary executable code, or an implementation utilizing GPU hardware, could improve the performance even further. Figure 6 shows the average execution time (with the respective standard deviation) based on a set of scale-space parameters. The horizontal axis denotes the number of used scales within the scale-space. The computational complexity is approximately linear in the number of scale levels in the scale-space. The deviation can be explained by the varying times consumed for fitting the Gaussian function. Our implementation utilizes a scale-space constructed by 46 scale levels.

III. SCALE-ADAPTIVE CLASSIFICATION

We define the certainty criterion κ based on the uncertainty measure of the scale estimation u, which corresponds to the standard deviation of the fitted Gaussian function:

$$\kappa(u) = \begin{cases} 1, & u < T_u \\ 0, & \text{otherwise} . \end{cases} \qquad (4)$$

For the threshold T_u, a value found to be reasonable during development was $T_u = 25$. If the scale response function ξ cannot be approximated using Gaussian fitting (either the

algorithm does not converge or the result is not in $\mathbb{R}$), the scale estimation is also defined as unreliable (setting $u = \infty$).

Consequently, we compute a distance matrix ($M = (m_{ij})$) containing the feature distances between each feature vector from the training set F_i and each feature vector from the evaluation set F_j. In a non-scale-adaptive classification scenario, the elements of M are defined as

$$m_{ij} = d(F_i, F_j) , \qquad (5)$$

where d is a distance measure. For scale-adaptive classification, we redefine the computation of the distance matrix as follows:

$$m_{ij} = \begin{cases} d(F_i, F_j), & \text{if } \kappa(u_j) = 0 \\ d(F_i, F_j), & \text{if } \kappa(u_i)\kappa(u_j) = 1 \\ & \wedge |\tilde{s}_i - \tilde{s}_j| < T_s \\ \infty, & \text{otherwise} , \end{cases} \qquad (6)$$

where $\tilde{s}_i$ and $\tilde{s}_j$ correspond to the estimated levels in scale-space. During experimentation the Euclidean distance measure is utilized as d. If the estimated scale of a feature vector in the evaluation set is unreliable ($\kappa(u) = 0$), the scale-adaptive classification is not feasible for this feature vector. We therefore utilize the standard classification scheme for such feature vectors. Feature distances for vectors in the evaluation set with a reliable scale estimation ($\kappa(u) = 1$), are only computed for reliable training vectors with an absolute relative scale difference below a certain threshold T_s. The remaining matrix elements are set to ∞, practically ignoring them for classification.

The threshold T_s adjusts the range of allowed relative scale differences. A small T_s leads to a narrow range of similar scales for classification, but also to a significantly reduced size of the training set. Higher values of T_s result in a diminished effect of scale adaptiveness due to the broad range of considered image scales. To avoid over-fitting, T_s for
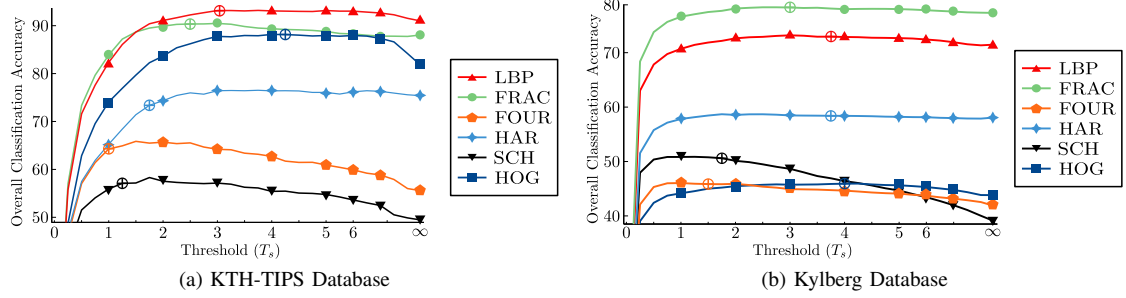
(a) KTH-TIPS Database

(b) Kylberg Database

Fig. 7: Classification Accuracies with varying Thresholds (T_s).

each distinct feature is found by exhaustively searching for the most appropriate value of T_s (with respect to the classification accuracy), using the opposing database. The thresholds used in Section IV-D for experiments based on the Kylberg database where found based on the KTH-TIPS database and vice versa. Figure 7 shows the classification accuracies with varying thresholds T_s. The $\oplus$ symbol denotes the actual thresholds used for our experiments discussed in Section IV-D. We observe the tendency, that the used thresholds are slightly too high in case of the Kylberg data set and too low in case of KTH-TIPS. We assume, that this is an effect caused by the big difference of the database sizes. A too low threshold would cause the training set to be at an unreasonable size, in case of the smaller KTH-TIPS. As a result, the identified threshold for the Kylberg data set is too high and vice versa for KTH-TIPS.

IV. EXPERIMENTS

We study the implicit scale selection effect of the k-nearest neighbor classifier to give a comprehensive overview of this effect in relation to the scale adaptive classification. Additionally, we compare the accuracies of the proposed scale-adaptive classification in comparison to the traditional classification based on the implicit scale selection.

A. Experimental Data

The experiments are based on two different gray-scale texture databases. The KTH-TIPS database [13] 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 [14], 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 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 patches per image) for experimentation. Table I provides detailed information of the databases used.

Database	KTH-TIPS	Kylberg
Texture Classes	10	28
Images per Scale	9	40
Different Scales	9	9
Max Relative Scale Difference	4.0	4.0
Images in Training Set	810	10080
Images in Evaluation Set	−	10080

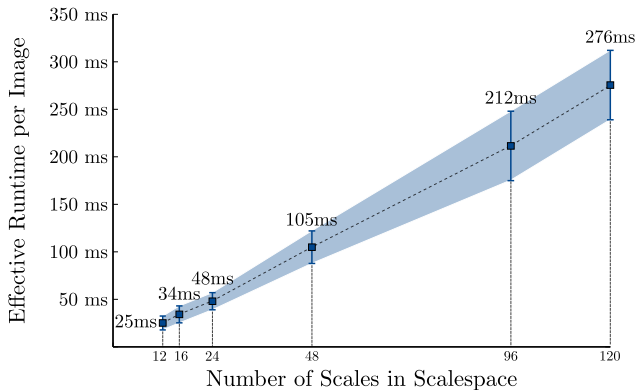TABLE I: Detailed Information on the used Texture Databases.

Fig. 6: The Effective Computational Time for Scale Estimation per Image.

B. Feature Extraction Methods

In order to measure the impact of scale-adaptive classification, we use the following texture features. For all features, the standard parameters are utilized, in order to avoid any bias.

- **Local Binary Patterns** [3] (LBP):
  LBP describes a texture by utilizing the joint distribution of pixel intensity differences represented by binary patterns. We employ LBP using the standard 8-neighborhood with a radius of 1 pixel.

- **Multi-Fractal Spectrum** [15] (FRAC):
  The local fractal dimension is computed for each pixel using three different types of measures for computing the local density. The feature vector is built by concatenation of these fractal dimensions. This method is by design scale-invariant. It is therefore interesting to
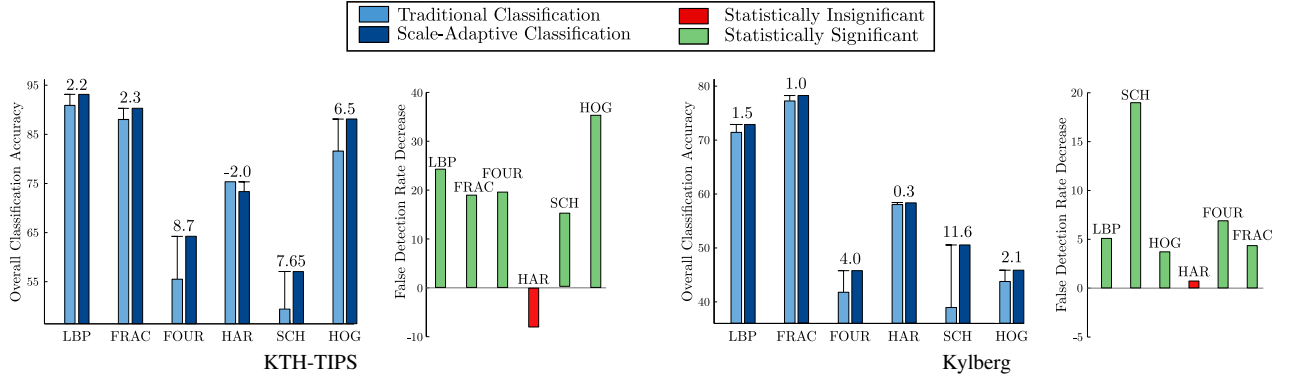
Fig. 8: Classification Results of the Scale-Adaptive Classification.

observe how the implicit scale selection and the scale-adaptive classification perform based on this method.

- **Fourier Ring Features** [16] (FOUR):
  After computing the Fourier power spectrum (low frequencies are aligned in the image center), the means of rings with a width of 1 pixel are calculated. For the experiments we concatenate the means of 17 equidistant rings between frequency 2 and 20 to construct a feature vector.

- **Haralick Contrast** [7] (HAR):
  The feature vector consists of the Haralick contrast feature [7] calculated for different offsets $(0, r)^T$, $(r, 0)^T$, $(r, r)^T$, $(r, -r)^T$. The final feature consists of the 16 dimensions (4 for each r, where $r \in \{1, 2, 3, 4\}$).

- **Shape Curvature Histogram** [9] (SCH):
  This feature represents the occurrences of the contour curvature values. The bin count which adjusts the granularity is set to 16.

- **Histogram of Gradients** [8] (HOG):
  The distribution of gradient orientations is used to describe a texture. This feature is used with the standard bin count of 9, which corresponds to an angular resolution of 20 degrees.

C. Experimental Setup

We perform a specific experiment to assess the implicit scale selection of the k-nearest neighbor classifier. For each database (Kylberg and KTH-TIPS), a distinct training and evaluation set (referred to as parent sets) consisting of images at all given scales is constructed. In order to assess the overall impact of scale differences on the implicit scale selection, we analyze the impact based on specific increasing relative scale differences. The relative scale difference is defined as the quotient of two scales. For each analyzed relative scale difference, we select specific data subsets containing two corresponding scales from the parent training and evaluation data, for classification. Each subset includes all images with two corresponding scales for a specific relative scale difference in a distinct training and evaluation set.

To evaluate the implicit scale selection of the k-nearest neighbor classifier, we assess four different properties, as

shown in Figure 4 and 5. The horizontal axes correspond to the relative scale difference used to create the specific data subsets. The vertical axes denote the classification accuracies averaged over the classification outcomes for each subset at a corresponding relative scale difference.

- **Probability of Equal Scales**
  To validate the assumption that an implicit scale selection is performed, we analyze scale distribution of the nearest neighbors. The set of nearest neighbors of all images for a specific relative scale difference form this distribution. In this case the vertical axis denotes the probability of the nearest neighbor of an image to be at the same scale as the evaluated image (corresponding to $k = 1$).

- **Implicit Scale Selection**
  For a relative scale difference, the standard k-nearest neighbor classification (with $k = 2$) is performed for each subset. The choice of $k = 2$ was motivated to enable comparability between the implicit and the restricted scale selection. This corresponds to the natural classification and reflects the effects of the implicit scale selection in comparison to the restricted scale selection.

- **Restricted Scale Selection**
  The restricted scale selection represents the assumed classification accuracy without implicit scale selection. We restrict the implicit scale selection by manually selecting the nearest neighbors for a 2-nearest neighbor classification for a specific image. One nearest neighbor corresponds to the feature vector with the minimal distance computed at the first scale in a specific subset, whereas the other nearest neighbor corresponds to the feature vector with the minimal distance computed at the other scale. This method suppresses the implicit scale selection of the classifier and allows a comparison between the restricted scenario and the natural classification, reflecting the effect of scale selection.

- **Optimal Scale Selection**
  The optimal scale selection denotes the upper bound achievable by selecting the same scales for training and evaluation.

Furthermore, we evaluate the proposed scale-adaptive classification based on the databases explained in Section IV-A utilizing the features in Section IV-B. We create two distinct sets for training and evaluation based on the Kylberg database using patches of image "A" for training and patches of image "B" for evaluation. Due do the restricted size of the KTH-TIPS database, we resorted to using leave-one-out cross validation in this specific experiment. The data sets are comprised of images at all available scales. For the final classification, we utilize the k-nearest-neighbor classifier. The accuracies for k reaching from 1 to 5 are averaged, in order to get more stable results.

D. Results

In Figure 4 and 5, the results of the experiments to assess the implicit scale selection of the k-nearest neighbor classifier are presented.

Based on the experimental results we observe a consistent beneficial effect of the implicit scale selection as compared to the restricted scale selection. We also observe that the probability of equal scales increases with scale difference. This indicates that the assumption that the intra-class variability within a scale of a texture generally is lower than the intra-class variability over multiple scales holds throughout the experiments. Interestingly this behavior is also observed for the FRAC method, which is supposed to be scale-invariant. This is an indication that feature vectors at different scales still exhibit a higher degree of difference as compared to feature vectors at the same scale. The results also show that features with higher classification accuracy in general benefit more from the implicit scale selection as compared to features with lower classification rates. We explain this by the more discriminative feature vectors which lead to larger differences between feature vectors at different scales, increasing the effect of the implicit scale selection. The gap between the optimal scale selection and the implicit scale selection suggests that there is room for improvement for a scale-adaptive classification.

Figure 8 presents the enhanced classification performance of the scale-adaptive classification in comparison to the traditional classification approach. We observe that the proposed method improved the classification accuracies with statistical significance (McNemar's test [17] with $\alpha = 0.05$) of all methods except for the HAR method. This holds for both independent data sets. An interesting fact is that the accuracy of the scale-invariant method FRAC could also be improved utilizing the scale-adaptive classification.

V. CONCLUSION

We have shown that the implicit scale selection mechanism of the k-nearest neighbor method has a significant impact on the classification accuracy of data sets comprised of various scales. We proposed a scale-adaptive classification framework to increase this beneficial effect utilizing a global scale estimation algorithm based on scale-normalized Laplacian of Gaussian extrema in scale space. Experimentation has shown that the classification accuracy can be statistically significantly improved by our method. Due to the partitioning performed by the k-nearest neighbor classifier we assume the effect of implicit scale selection to be more significant as compared to

other classification methods such as SVM. We therefore presume that the proposed framework might pose an even higher benefit in combination with such a classification method.

REFERENCES

[1] J. Zhang and T. Tan, "Brief review of invariant texture analysis methods," *Pattern Recognition*, vol. 35, no. 3, pp. 735–747, 2002.

[2] 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.

[3] 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.

[4] S. Liao, X. Zhu, Z. Lei, L. Zhang, and S. Li, "Learning multi-scale block local binary patterns for face recognition," in *Advances in Biometrics*. Springer, 2007, pp. 828–837.

[5] 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*, ser. Lecture Notes in Computer Science, vol. 4778, October 2007, pp. 168–182.

[6] 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.

[7] R. M. Haralick, Dinstein, and K. Shanmugam, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, pp. 610–621, November 1973.

[8] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1, 2005, pp. 886–893.

[9] M. Gadermayr, M. Liedlgruber, A. Uhl, and A. Vécsei, "Shape curvature histogram: A shape feature for celiac disease diagnosis," in *Medical Computer Vision. Large Data in Medical Imaging (Proceedings of the 3rd International MICCAI - MCV Workshop 2013)*, ser. Springer LNCS, vol. 8331, 2014, pp. 175–184.

[10] S. Hegenbart and A. Uhl, "A scale-adaptive extension to methods based on lbp using scale-normalized laplacian of gaussian extrema in scale-space," in *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing ICASSP '14*, 2014, pp. 4352–4356.

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

[12] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91–110, November 2004.

[13] E. Hayman, B. Caputo, M. Fritz, and J.-O. Eklundh, "On the significance of real-world conditions for material classification," in *Proceedings of the European Conference on Computer Vision*, vol. 4, 2004, pp. 253–266.

[14] G. Kylberg, "The kylberg texture dataset v. 1.0," Centre for Image Analysis, Swedish University of Agricultural Sciences and Uppsala University, Uppsala, Sweden, External report (Blue series) 35, September 2011. [Online]. Available: http://www.cb.uu.se/~gustaf/texture/

[15] Y. Xu, H. Ji, and C. Fermüller, "Viewpoint invariant texture description using fractal analysis," *International Journal of Computer Vision*, vol. 83, no. 1, pp. 85–100, 2009.

[16] M. Häfner, L. Brunauer, H. Payer, R. Resch, F. Wrba, A. Gangl, A. Vécsei, and A. Uhl, "Pit pattern classification of zoom-endoscopic colon images using DCT and FFT," in *Proceedings of the IEEE International Symposium on Computer-Based Medical Systems (CBMS'07)*. Maribor, Slovenia: IEEE Computer Society CPS, Jun. 2007, pp. 159–164.

[17] Q. McNemar, "Note on the sampling error of the difference between correlated proportions of percentages," *Psychometrika*, vol. 12, no. 2, pp. 153–157, June 1947.