Texture-Based Features for Clothing Classification via Graph-Based Representation

Srisupang Thewsuwan and Keiichi Horio

Graduate School of Life Science and Systems Engineering
Kyushu Institute of Technology
2-4 Hibikino, Wakamatsu-ku, Kitakyushu-shi, Fukuoka 808-0196, Japan
E-mail: Thewsuwan-srisupang@edu.brain.kyutech.ac.jp, horio@brain.kyutech.ac.jp

Abstract This paper proposes texture-based features for clothing category classification based on graph representation. Recently, graph-based representation has been used for texture characterization to aid texture analysis. In this work, graph-based theory is applied to characterize the local image structure. The rotation invariance uniformity (riu2) of local binary pattern mapping is adopted to represent feature descriptors. The proposed approach is evaluated by using the Brodatz and UIUC standard texture databases, and a clothing dataset. The proposed method is shown to be more effective for clothing classification than other methods.

Keywords: texture classification, graph-based representation, local binary pattern mapping

1. Introduction

Clothes can be defined as deformable and non-rigid objects which are difficult to classify when they are crumpled in a pile of laundry. This task is extremely challenging because clothing in a free configuration can be highly wrinkled, tangled and in a huge variety of poses. Therefore, it is difficult to encode a clothing category into a generic visual representation. Figure 1 shows an example of clothing with a textured surface lying on a table in various poses.

To meet the above challenge, an efficient feature descriptor is still required in this research area. Appearance information of clothes images has been used to develop feature descriptors for clothing classification [1]-[4]. For instance, Yamazaki and co-workers [1], [2] have proposed clothing classification system that uses fabrics, wrinkles and cloth overlaps as features combining global and local information to derive input information [3], [4]. The results of combining a local feature with a global characteristic achieved high classification performance. Accordingly, the wrinkles of fabric can be used as visual information directly related to the object's surface. These works inspired us to develop a feature descriptor for seeking unique features on the surface of fabrics, for example, texture patterns. However, it remains challenging to investigate the feature-enriched representation of clothes, because the natural texture characterization of fabrics can be presented as random and persistent stochastic patterns [5].

Texture is an efficient way to represent the appearance of an object in an image [6]. It can provide important information for object identification based on a physical characteristic. Using such information, texture analysis research has greatly advanced, resulting in enhanced texture pattern descriptors. In recent years, image representation by graph theory has been employed for texture analysis [7]. A graph-based representation [8] can express the context...
surrounding each pixel and the relation among structural texture elements, which are crucial feature properties used to distinguish different classes of images. Although the graph-based representation is effective for texture analysis and classification, there are various challenging issues that should be investigated before it can be applied to clothing category classification.

This paper proposes texture-based features for clothing category classification based on a graph representation. The advantage of the rotation uniformity of local binary pattern (LBP) mapping and graph theory is adopted to construct the feature descriptor. Therefore, the empirical synergy between LBP mapping and graph theory is a promising direction in this work. Graph-based theory [8] is applied to represent the spatial relation between image pixels and their neighbors. To increase the effectiveness, the rotation invariance uniformity of the riu2 mapping [9] is employed to encode the spatial arrangement of the local spatial distribution. A clothing database and the standard Brodatz [10] and UIUC [11] texture databases are used for evaluation. The experimental results show the effectiveness of the proposed method compared with texture analysis based on conventional methods in terms of the success rate of classification.

To describe the proposed method, the remainder of the paper is organized as follows. Section 2 explains our proposed method, which includes the main topics of graph-based pixel representation and the binary pattern transformation process. The feature descriptors are explained in Sect. 3, followed by experimental results and discussion in Sect. 4. Finally, Sect. 5 concludes this work.

2. Proposed Method

This section includes the graph-based pixel representation and binary pattern transformation. Figure 2 shows an overview of our proposed approach for clothing classification. The system can be separated into two parts. The first part illustrates the process of graph-based pixel representation including a deterministic weighted graph and the radial distance pattern mapping. The weighted graph is defined by the local grayscale difference. The radial distance mapping is
constructed using the Euclidean distance between a node and a pairwise connection. We set radial distance $r_{\text{max}}$ equal to 1, 2, and 3 as a fixed window mapping with a multi-scale neighborhood. The second part is the binary pattern transformation. The deterministic weighted graph is applied to extract local textural information. The extraction process includes weighted binary thresholding and weighted binary non-thresholding. Feature descriptors are derived by using the $\text{riu}^2$ mapping technique, in which the features of the final histogram are concatenated histograms obtained from the weighted binary thresholding and the weighted binary non-thresholding.

2.1 Graph-based pixel representation

2.1.1 Weighted graph

Weighted edges are used to represent a structure in which pairwise connections have numerical values. The simple data structure of an image consists of pixel information including color values and coordinates. The weight of the graph is defined by different pixel intensities, that is, co-occurring pixels with different intensities can be used to obtain the weights of edges, and consequently this approach can characterize local textural information. The weight of an edge $W(e)$ is given by

$$W(e) = \left\{ \begin{array}{ll}
\sqrt{(i-i')^2 + (j-j')^2} \leq r_1 \\
0 & \text{otherwise}
\end{array} \right. \quad (1)$$

Based on the radial distance $r$, the three radial patterns are denoted by $P(1,4)$, $P(2,8)$ and $P(3,16)$ (see Fig.2 for an example).

2.1.2 Radial distance pattern mapping

The graph structure is the main idea for image representation to reflect the structure of an input image. We propose radial distance pattern mapping for multiple-scale analysis by increasing the scale and the resolution of the pixel connectivity. A radial graph can be expanded to incorporate other vertices based on the Euclidean radial distance. This is the simplest way to represent information about an image on a per-pixel basis. A pixel of the texture image is denoted as a vertex $(v_{ij})$ in the graph. Two vertices are connected when the Euclidean distance between them is less than or equal to the value $r_{\text{max}}$. The radial distance pattern mapping is given by

$$e = (v_{ij}, v_{i'j'})$$

$$\epsilon \in l \times l \| l \| \left\{ \begin{array}{ll}
\sqrt{(i-i')^2 + (j-j')^2} \leq r_1 \\
r_1 < \sqrt{(i-i')^2 + (j-j')^2} \leq r_2 \\
r_2 < \sqrt{(i-i')^2 + (j-j')^2} \leq r_3 \\
\vdots \\
r_{\text{max}-1} < \sqrt{(i-i')^2 + (j-j')^2} \leq r_{\text{max}}
\end{array} \right. \quad (2)$$

In this work, we set $r_{\text{max}} = 1, 2,$ and $3$ to generate three patterns mapping. These patterns consist of different considering neighborhoods or dimensions which are equal to 4, 8, and 16, respectively, as indicated by gray color in the graph-based pixel representation in Fig.2.

2.2 Binary pattern transformation

This work considers the edges as vectors that have magnitude and direction properties. The magnitude can be referred as non-direct edges in a graph, whereas the direction defines direct edges in the graph. The magnitude of the weighted graph is obtained as an absolute value of the weighted graph. The sign of the weighted graph value is used to determine the direction of an edge. The binary pattern transformation can be generated by thresholding. For the magnitude value, this approach requires a threshold to generate the binary pattern, whereas the sign can present by itself. Accordingly, weighted binary thresholding and a weighted binary non-thresholding are denoted as the magnitude and the direction of edges in this paper. These two approaches can extract different local textural information essential for texture analysis.

2.2.1 Weighted binary thresholding ($WB^t$)

$$WB^t(e) = \left\{ \begin{array}{ll}
1 & |W(e)| \leq t^A \\
0 & \text{otherwise}
\end{array} \right. \quad (3)$$

where $t^A$ is adaptive local threshold as below. $t^A$ is generated by averaging the absolute values of the weighted edges in the graph in Eq. (2). This approach is performed by setting the pixels whose weights are less than or equal to threshold $t^A$ to 1, while the remaining pixels are set to 0. A set of binary patterns $WB^t$ used in multiple-scale analysis $r$ is defined as follows:
where $p_1$, $p_2$, and $p_3$ are equal to 4, 8, and 16, respectively. The local feature descriptor based on the $WB_t$ property is given by

$$\Phi(v_i) = [WB_{t1}, WB_{t2}, WB_{t3}]$$  \hspace{1cm} (5)$$

2.2.1 Weighted binary non-thresholding ($WB^{nt}$)

$$WB^{nt}(e) = \begin{cases} 1 & \text{sign}(W(e)) > 0 \\ 0 & \text{otherwise} \end{cases}$$  \hspace{1cm} (6)$$

A set of binary patterns $WB^u$ used in multiple-scale analysis $r$ is defined as follows:

$$WB^{nt}(e_{v_i,v_j}) = \begin{bmatrix} WB_{r1}^{nt}(e_{v_i,v_1}), \cdots, WB_{r1}^{nt}(e_{v_i,v_{p1}}) \\ WB_{r2}^{nt}(e_{v_i,v_1}), \cdots, WB_{r1}^{nt}(e_{v_i,v_{p2}}) \\ WB_{r3}^{nt}(e_{v_i,v_1}), \cdots, WB_{r1}^{nt}(e_{v_i,v_{p3}}) \end{bmatrix}$$  \hspace{1cm} (7)$$

where $p_1$, $p_2$, and $p_3$ are equal to 4, 8, and 16, respectively. The local feature descriptor based on the $WB^{nt}$ property is given by

$$\Psi(v_i) = [WB_{r1}^{nt}, WB_{r2}^{nt}, WB_{r3}^{nt}]$$  \hspace{1cm} (8)$$

3. Feature Descriptors

After obtaining the binary patterns, the $riu2$ mapping is employed to consider the uniform binary patterns. The uniform LBP mapping was proposed as means of improving the rotation invariance of image textures. The $riu2$ mapping is used to encode the uniformity of the primitive texture when the binary pattern contains at most two transitions ($U \leq 2$) between 0 and 1. Suppose that $U$ defines a bit-wise transition, where the binary pattern of 00000000 has a $U$ value of 0, whereas the binary pattern of 00000110 has a $U$ values of 2, which means that this binary pattern is uniform pattern. In another case, if the binary pattern of 00110110 pattern is non-uniform. This step enables us to analyze the uniformity pattern of vertices to ensure rotation invariance. In practice, this step is implemented by using a look-up table of $2^v$

Fig. 3 Example images from our clothing dataset, which includes towels, pants, skirts, and shirts

4. Results and Discussion

In the following experiments, the nearest neighborhood with the Euclidean distance is used as a discrimination function, with 10-fold cross-validation used to evaluate discrimination performance. This simple classifier of the nearest neighborhood was chosen rather than a more sophisticated classifier to demonstrate the importance of the features in the classification task. The experimental results are separated into the results of the proposed approach with the clothing dataset and with texture databases.

4.1 Databases

In the experiments, two standard texture databases and a clothing dataset are used for evaluation as follows:

[Clothing dataset] The sample images were captured by using an Asus Xtion PRO camera with a controlled environment, for example, controlled illumination and scale. The resolution of each image was 640×480 pixels. The scale distance between the camera and the clothing was 900 [mm]. The databases include four categories: towels, shirts, pants, and skirts. Each category consists of five styles of clothing (see Fig.3 for examples). This dataset contains 1,000 images, including 50 images per style, 250 images per category. All images were captured by throwing each piece of...
clothing randomly onto the surface. Variations in the pose and deformation of the textured surface, such as crumpled and smooth clothes make it challenging to classify such items in a database. Moreover, the clothes dataset has large intraclass variations. Note that color information was not used in the following experiment; i.e. grayscale images were used.

[Brodatz texture album] This database is used in texture analysis and is a benchmark for evaluating methods. Images included in the database are arranged in 100 classes, each class containing 10 grayscale samples of $128 \times 128$ pixels obtained by splitting the data of each class into 10 non-overlapping sub-images.

[UIUC database] This database includes the images which is difficult to be correctly classified. The images have significantly different viewpoints and scales due to perspective distortion and non-rigid transformation. The image size is $128 \times 128$ pixels. For each of the 25 classes, 40 grayscale images were considered in the experiments.

4.2 Results of using proposed feature descriptor with clothing dataset

Figure 4 illustrates examples of the final histograms obtained using the proposed method from the clothing dataset. The experimental results are shown as confusion matrices using the proposed feature and other features for comparison including the feature descriptor [12], Gabor filter [13], LBP, LBP$_{iu}$ [9], and CLBP [14] as illustrated in Fig. 5. In this figure, diagonal values indicate the success rate of each clothing class. Rows correspond to the true class, and columns represent the predicted class. The proposed features achieve a relatively stable performance among all categories or classes. Using the proposed feature, the highest success rate of 76.08% was achieved for the class of pants, while the other methods scored success rates of less than 67%. The worst result was obtained by the LBP feature and had a success rate of 39.72%. Note that shirts and pants are more susceptible to being textured and crumpled than towels and shirts. This likely leads to higher inter-class similarities.

The weighted binary thresholding and non-
Thresholding can be used to extract local textural information of different properties. Note that combining \((WB_t, WB_{nt})\) properties in a deterministic weighted graph results in better performance than separating \(WB_t\) or \(WB_{nt}\). To achieve high local discrimination capability, the local grayscale difference discriminative information in terms of \(WB_t\) and the local textural distribution by adopting rotation invariant microstructure in term of \(WB_{nt}\), were proposed to distinguish different local structures information in this paper. Although the clothing dataset has high intraclass variation, the experimental results for the proposed feature show that it is more effective for extracting local discriminative information as texture-based features for clothing classification than other feature methods.

### 4.3 Results of using proposed feature descriptor with texture databases

Table 1 lists the success rate \([\%]\) of the proposed approach and other texture analysis methods using the Brodatz and UIUC databases and the clothing dataset. Our proposed method achieved the highest success rate of 90.92\% with Brodatz texture dataset. This result shows that the combined local grayscale distribution and the second-order local structure information in the deterministic weighted graph has very high local discrimination capability. In the UIUC challenging database, the experimental results show that the CLBP method achieved the best success rate of 93.64\%, following the proposed method with a success rate of 87.92\%. From the numbers of features listed in Table 1, we can see that the CLBP can extract more discriminative information than the proposed method. Moreover, the LBP\textsuperscript{riu}\textsuperscript{2} can also achieve good result which is comparable with the proposed method. This experimental result noticed us that the UIUC database is affected by the local grayscale different information. On the other hand, the proposed method achieved a success rate of almost 88\%, which is comparable with other methods. Accordingly, these experimental results confirmed that the proposed feature descriptor is effective for describing the texture pattern on images and for clothing category classification.

### 5. Conclusions

In this paper, we proposed a new feature descriptor for clothing category classification. The synergy between graph theory and LBP mapping was used as an approach for characterizing the texture pattern on images. The experimental results show that the proposed method can obtain good performance for clothing classification compared with other methods.

### References


Srisupang Thewsuwan received her M.E. degree from King Mongkut’s University of Technology Thonburi (KMUTT) in 2015, and Ph.D. degree from Kyushu Institute of Technology in 2018. Her current research interests are computer vision and pattern recognition.

Keiichi Horio received his M.E. and Ph.D. degrees from Kyushu Institute of Technology in 1998 and 2001, respectively. He is an Associate Professor at the Graduate School of Life Science and Systems Engineering, Kyushu Institute of Technology. His research interests include image processing and data analysis.

(Received May 14, 2018; revised October 26, 2018)