Region Description Using Extended Local Ternary Patterns

Wen-Hung Liao

Dept. of Computer Science, National Chengchi University, Taipei, Taiwan
whliao@cs.nccu.edu.tw

Abstract

The local binary pattern (LBP) operator is a computationally efficient local texture descriptor and has found many useful applications. However, its sensitivity to noise and the high dimensionality of histogram associated with a mediocre size neighborhood have raised some concerns. In this paper, we attempt to improve the original LBP by proposing a novel extension named extended local ternary pattern (ELTP). We will investigate the characteristics of ELTP in terms of noise sensitivity, discriminability and computational efficiency. Preliminary experimental results have shown better efficacy of ELTP over the original LBP.

1. Introduction

Local binary pattern is a computationally efficient local texture descriptor that has been applied successfully to tasks such as texture classification, face recognition, and background modeling [1]. However, there exist several limitations of LBP that hinder its capability in certain situations. For example, LBP is quite sensitive to random noise in near-uniform image regions. Moreover, LBP with mediocre size of sample points will produce a feature representation with very high dimension. The former issue is, to some extent, resolved by the introduction of local ternary patterns [2]. Yet LTP has a much larger histogram size than the original LBP. The latter problem is generally settled by merging or grouping patterns to reduce the size of the histogram. Yet the dimensionality reduction process usually casts a negative effect on LBP's capability to accurately describe a region. So far, no proposed solution can address both issues simultaneously and effectively.

The objective of this research is to improve the original LBP using a novel extension named extended local ternary pattern. The proposed ELTP accomplishes better tolerance to noise through the incorporation of ternary representation, while at the same time controls the histogram size by merging patterns using a distance measure defined over the ternary digit pattern space.

The rest of this paper is organized as follows. In Section 2 we briefly reviewed previous work, focusing on those directly related to local ternary patterns. Section 3 describes the formulation of extended local ternary patterns, along with some possible variations. Basic properties of the newly proposed ELTPs are also discussed. Section 4 presents some preliminary experimental results and comparative analysis. Section 5 concludes this paper with a conclusion and outlook on future work.

2. Related work

There are various extensions and modifications of the original LBP following its first introduction by Ojala *et al*. [3]. A good source of references can be found in [4]. Since our investigation focuses on issues regarding noise sensitivity and histogram bin size, we will restrict our discussion of related work to these subjects.

According to the original definition of LBP, pattern with a mediocre size of sampling points (P in LBP(P,R)) will generate a histogram of rather high dimensionality. For example, LBP(16,2) will generate a histogram of size 2^{16}=65536, which is not suitable for region description. (A 32x32 image patch will have at most 1024 distinct patterns, resulting in a very sparse representation.) To address this issue, Mäenpää *et al.* [5] proposed two approaches to select a subset of LBP for texture classification. The first method starts with a single pattern and iteratively expands the

pattern collection using a training set. However, the patterns thus chosen tend to depend on the training data employed. The second approach reduces the histogram size by dividing the patterns into uniform and non-uniform ones. The size can be further reduced to *P+2* (*P* is the number of sample points) by incorporating rotation-invariance. Such modifications may prove effective for texture disambiguation, but may face difficulties for tasks such as face or object recognition [6].

The second critical issue regarding LBP is its sensitivity to noise. Let us examine the 3x3 image in Fig. 1. The corresponding LBP is 11001011, or 203. If we change the intensity value of the center-left pixel from 54 to 53, we will obtain a different LBP: 01001011, or 75 (shown in Fig. 2). Notice that these two bit patterns are still quite similar, with their Hamming distance equal to 1. But the values become distinct when converting into decimal representation.
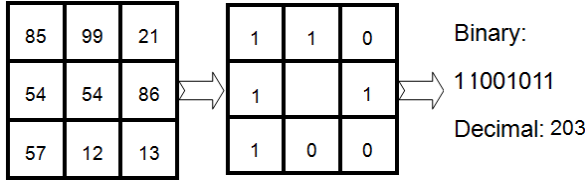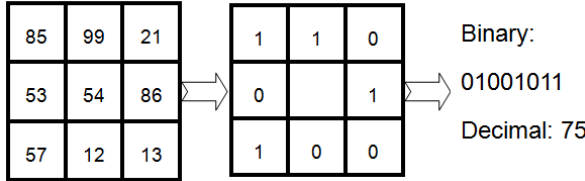

Fig. 1 Calculation of the LBP.

Fig.2 LBP obtained by modifying the value of the center-left pixel from 53 to 53.

.There exist several methods to compute the distance between two histograms. But either histogram intersection (Eq.1) or χ^2 distance (Eq. 2) considers individual bins separately.

$$HI(\boldsymbol{S},\boldsymbol{T}) = \sum_i \min\left\{\frac{S(i)}{\sum S(j)}, \frac{T(i)}{\sum T(j)}\right\} \quad (1)$$

$$\chi^2(\boldsymbol{S},\boldsymbol{T}) = \sum_i \frac{(S(i)-T(i))^2}{S(i)+T(i)} \quad (2)$$

As a result, the slight perturbation caused by replacing a single pixel value yields a rather significant change in pattern distribution and distance measure.

Local ternary pattern seems to be a natural extension of the original LBP to deal with this problem. In [2], Tan *et al.* proposed to use a base-3 pattern to represent the region. The LTP can be calculated according to Eq. (3):

$$LTP(i) = \begin{cases} 1 & \text{if } P(i) - P(0) > \theta \\ 0 & \text{if } |P(i) - P(0)| \le \theta \\ -1 & \text{if } P(i) - P(0) < -\theta \end{cases} \quad (3)$$

where *P(0)* is the intensity of the center pixel, and θ is a pre-defined threshold. Using this new representation, the two image patches in Figs. 1 and 2 can both be converted to the same ternary pattern shown in Fig.3.
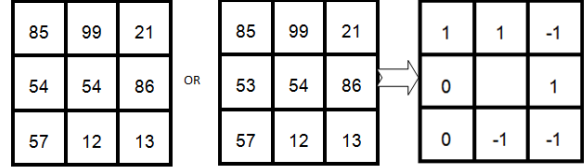

Fig. 3 LTP obtained by setting θ=5.

While this approach seems to have addressed the noise sensitivity issue, it actually creates another problem regarding the histogram dimensionality. For example, *LTP(8,1)* will generate a histogram of size 3^8=6551. *LTP(16,2)* will create a histogram of size 3^{16}=43,046,72, which is deemed inappropriate for practical implementation.

Dimensionality reduction is achieved in [2] by decomposing the LTP into two LBPs (upper and lower), as demonstrated in Fig. 4. The distance between two LTP is calculated by combing the results from upper and lower LBP, respectively.
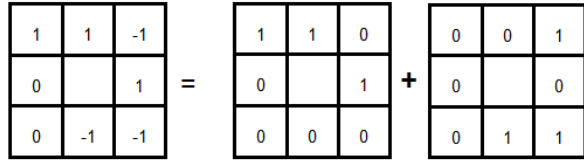

Fig. 4 Decomposing LTP into two LBPs.

The above coding scheme has brought relief to the histogram dimensionality problem. However, it also adversely affects the pattern's tolerance for noise, which is considered to be the key strength of LTP. According to our experiments, LTP-UL-LBP performs even worse than the original LBP in the presence of noise. An effective coding scheme needs to be developed to maintain compactness of the feature vector, while at the same time retain its ability to faithfully describe the object of interest in a noisy environment.

3. Extended Local Ternary Patterns

There are two conflicting factors that affect the performance of LBP/LTP. On one hand, when we

employ a larger number of sample points or use a base-3 representation, we will achieve better feature resolution. One the other hand, too fine the resolution implies sensitivity to minor changes in the pattern, as well as difficulties in actual implementation. (Consider the case of LTP(16,2)). The proposed ELTP attempts to strike a balance by using a clustering method to group the patterns in a meaningful way, as explained in the following.

3.1. Basic formulation

The process for converting a region into its ELTP representation (Eq. (4)) is very similar to that described in [2], i.e.,

$$LTP(i) = \begin{cases} 1 & \text{if } P(i) - P(0) > \theta \\ 0 & \text{if } |P(i) - P(0)| \leq \theta \\ -1 & \text{if } P(i) - P(0) < -\theta \end{cases} \quad (4)$$

Instead of employing a fixed threshold θ, however, we propose to assign its value based on the local statistics of the pattern. Specifically, we will use Eq. (5) to compute θ:

$$\theta = \alpha \times \sigma \quad (0 < \alpha \leq 1) \quad (5)$$

where σ is the standard deviation of the local patch, and α is a scaling factor. Such a formulation helps to retain one favorable property of LBP: invariance with respect to illumination transformation, as illustrated in the following example

Fig. 5 depicts the LTP of two image patches using our proposed criteria (α=0.3). The intensity value of the right patch is obtained via a simple linear transform:

$$R(i) = L(i) \times 3 + 10 \quad (6)$$

10	40	45
23	25	23
40	15	35

or

40	130	145
79	85	79
130	55	115

➡
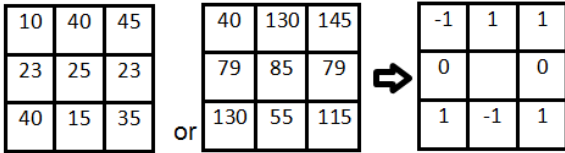
-1	1	1
0		0
1	-1	1

Fig. 5 Invariance of ELTP under gray-level transformation.

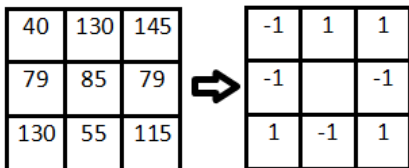If we use fixed threshold, say θ=5, the right region will have a different LTP, as shown in Fig. 6.

40	130	145
79	85	79
130	55	115

➡

-1	1	1
-1		-1
1	-1	1

Fig. 6 LTP using a fixed threshold (θ=5)

3.2. Dimensionality reduction

As discussed previously, using a base-3 system for representing feature patterns will increase the feature dimension in a drastic manner. It is therefore necessary cut down the size of the histogram by grouping patterns. But how does one achieve this goal in a sensible manner? Here we propose to form the groups based on pattern similarity. Suppose $x = (xn-1,...,x0)$ and $y=(yn-1,...,y0)$ are two ELTP strings, the distance between x and y can be calculated using their Hamming distance:

$$D(x, y) = \sum_{i=0}^{n-1} |x_i - y_i| \quad (6)$$

For a ternary string of length n,

$$\max \quad D(x, y) = 2n \quad (7)$$

The similarity (or affinity) between two ELTP strings can therefore be defined as:

$$A(x, y) = 1 - \frac{D(x,y)}{2n} \quad (8)$$

When there is a need to group patterns, those with larger affinity should be merged together. Specifically, if P is the number of sample points, there will be at most 3^P distinct ELTP strings. To reduce the size of the feature dimension from 3^P to K, we will first compute the similarity between any two ELTP strings to form a $3^P \times 3^P$ symmetric affinity matrix. This will transform the original dimensionality reduction problem into a graph partitioning problem, which can be solved using spectral clustering algorithms [7]. It should be noted that the same process can be applied to reduce the feature dimension of LBP. The only difference lies in the way one computes the similarity measure (Eq. 9):

$$A_2(x, y) = 1 - \frac{D(x,y)}{n} \quad (9)$$

Fig. 7 summarizes the procedure for reducing feature dimension for the proposed ELTP representation.

S1. Choose P(sample points) and K(histogram size)
S2. Form a $3^P \times 3^P$ affinity matrix using Eq. (8)
S3. Perform a K-way partition of the 3^P patterns using spectral clustering.
S4. Merge those patterns belonging to the same partition into a single bin in the histogram.
S5. Use the K-dimensional histogram for feature representation.

Fig. 7 Dimensionality reduction process for ELTP.

Following the K-way partitioning, it is possible to compute the mean distance of patterns belonging to the same cluster. The mean distance is regarded as an indicator of cluster homogeneity, and can be used to assign the weight w_i in weighted χ^2 distance (Eq. 10). Clusters of greater homogeneity should get more weight.

$$\chi_w^2(\boldsymbol{S}, \boldsymbol{T}) = \sum_i w_i \frac{(S(i) - T(i))^2}{S(i) + T(i)} \qquad (10)$$

The uniform pattern argument also applies in the ternary case, with a slight modification in the definition. If we consider only the number of transitions in the ternary digit pattern, then both 12222111($1 \rightarrow 2$, $2 \rightarrow 1$) and 02220000 ($0 \rightarrow 2$, $2 \rightarrow 0$) will be uniform patterns. However, if we take into account the amount of transitions, then the former pattern (transition amount=2) will be more uniform than the latter (transition amount=4).

Finally, ration-invariance version of the ELTP can be obtained by defining a new distance measure (Eq. 11):

$$D'(x, y) = \min_{j=0\ldots n-1} D(x, ROT_j(y)) \qquad (11)$$

where ROT_j denotes the circular shift operation by j digits.

4. Experimental Results

We present preliminary experimental results comparing the performance of the original LPB, LTP-UL-LBP and the newly proposed ELTP in term of noise sensitivity. For the purpose of comparison, we set $P=8$ and $K=256$. The scaling factor α for defining ELTP is set to 0.3 in the experiment. We use lena image to perform the test. The image is corrupted with Gaussian noise of different scales. Image patches are randomly selected from the noisy image and the corresponding LBP, LTP and ELTP are calculated. To evaluate noise immunity, we compute the histogram intersection between the original patterns and their noisy counterparts. The results are depicted in Fig. 8.

Generally speaking, ELTP is least sensitive to perturbations, especially at high noise levels, using the same size of feature vector ($K=256$). The original LTP was designed to have better noise resistivity, yet the coding scheme (by decomposing into upper and lower LBP) counters all the benefits. As for the computational complexity, the K-way partition needs to be performed only once. After that, the grouping of patterns can be done using a fairly simple table-lookup method.
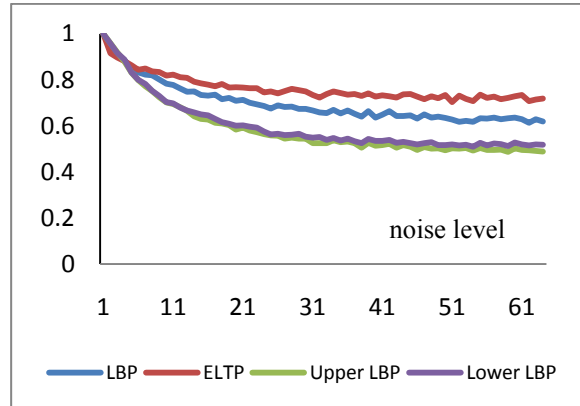

Fig. 8. Performance comparison.

5. Conclusions

A novel scheme of defining local ternary patterns and a systematic approach for grouping these patterns have been devised in this paper. Preliminary experimental analysis showed encouraging results using the proposed ELTP for region description.

Future work includes an in-depth investigation of different spectral clustering algorithms and how they affect the partitioning results. More importantly, we will examine the efficacy of the proposed ELTP to machine vision applications such as texture classification, face or facial expression recognition and background modeling.

References

[1] T. Mäenpää and M. Pietikäinen, Texture Analysis with Local Binary Patterns. In: Chen CH & Wang PSP (eds) *Handbook of Pattern Recognition and Computer Vision*, 3rd ed, World Scientific, pp. 197-216, 2005.

[2] X. Tan and B. Triggs. Enhanced Local Texture Feature Sets for Face Recognition Under Difficult Lighting Conditions". In *Analysis and Modeling of Faces and Gestures*, volume 4778 of LNCS, pp.168–182. Springer, 2007.

[3] T. Ojala, M. Pietikäinen, and D. Harwood, A Comparative Study of Texture Measures with Classification Based on Feature Distributions, *Pattern Recognition,* vol. 29, pp. 51-59,1996.

[4] http://www.ee.oulu.fi/mvg/page/lbp_bibliography

[5] T. Mäenpää, T. Ojala, M. Pietikäinen and M. Soriano, Robust Texture Classification by Subsets of Local Binary Patterns. *Proc. 15th International Conference on Pattern Recognition,* Vol. 3:pp. 947-950, 2000.

[6] T. Ojala, M. Pietikainen, and T. Maenpaa, Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns, *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 24, no. 7, pp. 971-987, July 2002

[7] U. von Luxburg, A Tutorial on Spectral Clustering. *Statistics and Computing*, Vol. 17(4) pp. 395-416, 2007.