AGE CLASSIFICATION USING ISOMETRIC AND ANISOMETRIC CROSS DIAGONAL -LOCAL DIRECTION TERNARY MATRIX

AKARAPU SWARNA¹, V. VENKATA KRISHNA², SUPREETHI K.P³

¹Research scholar, JNTUH, Asst professor, Department of Computer Science Engineering, VJIT, Hyderabad, Telangana, India,
²Professor, Department of Information Technology, Vasavi College of Engineering, Hyderabad, Telangana, India,
³Professor, Department of CSE, JNTUH College of Engineering, Hyderabad, Telangana, India,

E-mail: swarnaa@vjit.ac.in, swarnaa@vjit.ac.in, drvvk@staff.cee.ac.in, supreethi.pujari@jntuh.ac.in

ABSTRACT

Aging has a massive effect on the features and appearance of the human face. Even though various traits are utilized to estimate human age, this article focuses on age classification using prominent texture and edge-based feature vectors. Most of the local-based methods derive features on a circular window. The applications related to facial skin require the computation of elliptical or anisotropic features since the lips, eyes, and other prominent facial skin features primarily represent elliptical shapes. Further, the locally-based approaches are prevalent in estimating human age; however, most of these methods are intensity-based and sensitive to noise illumination changes, thus may not provide better results. To address this, the local directional patterns (LDP) are proposed, which derives features based on the top edge responses in all eight directions in the form of binary patterns of a 3x3 window. The disadvantage of LDP is finding the threshold for top edge responses. This paper derived an automatic process for deriving thresholds and explored the derivation of the ternary pattern instead of binary patterns. To reduce the complexity, the 3x3 window is divided into cross diagonal -local direction Ternary matrix (CD-LDTM) on both isometric (ICD-LDTM) and anisometric (ACD-LDTM) local structures. The facial features derived by the proposed ICD-LDTM and ACD-LDTM descriptors are fed to machine learning classifiers for age classification purposes. The experimental results demonstrate that the proposed strategy is effective.

Keywords: Edge Responses; Intensity; Elliptical Features; Machine Learning

1. INTRODUCTION

The accessibility of many images has resulted in image sharing via internet technology. In recent years the rapid growth and improvements in technology, a wide range of electronic devices, such as smartphones, surveillance and digital cameras are now available that can produce high-quality photographs and movies. These devices have virtually become an inseparable part of our lives. The high-resolution digital images and films gathered from them, causing the size of visual data to rise exponentially. Our brains can grasp images/videos faster than text; they can make communication easier. That is why the image is used as the primary source of information in many uses such as computer vision, medical imaging, biometrics, object recognition, facial expression analysis, age classification, and so on. Because human annotation is a hard and time-consuming procedure, an automated and intelligent solution that can classify or segregate similar visual material from the digital database is required. This segregation of relevant or the similar class of images automatically in to one group is the basic aim of a classification problem [1-9].

2. LITERATURE SURVEY

As essential visual signals, human faces convey a substantial quantity of nonverbal information to aid human-to-human communication in the real world. Identification, emotion, age, gender, and race are just a few of the personal qualities found on the human face. The information conveyed by the face has piqued the interest of researchers in the field of age classification, face recognition, and facial
The local-based approaches on a 3x3 window have become popular for the last few years. Texture analysis is a well-defined technique that provides the image's surface information, such as roughness, homogeneity, irregularity, and so on. The significant information or content extracted from the facial images is treated as the feature vector. The derivation process of feature vector from the facial skin texture is the most crucial factor for determining the efficacy of various applications related to facial skin. The derivation of the feature descriptor to extract the features of a face, on the other hand, is complicated. The fundamental drawback of the age classification methods is that there is no unique way of deriving feature vectors for all facial images. The facial image's properties change when the same scene is taken in a different lighting situation and from a different viewing point. As a result, succinct and unique derivation is not possible for the same set of images. The key problem for age classification techniques is feature extraction, from facial images. The texture provides the image's surface information, such as roughness, homogeneity, irregularity, and so on. Texture analysis is a well-defined technique that may be used on various classification approaches [19,20,43]. This paper has done a complete literature survey on the popular local based approaches and its variants and found the following: The Local Pattern descriptors effectively extract significant textural properties from a given image based on the region of interest, resulting in a higher classification rate. That is why image feature extraction has been paying more attention to local pattern descriptors. The local binary pattern (LBP) is one of the prominent local descriptors [21]. The LBP characterises an image's textural feature by estimating the relationship between the centre pixel and its local neighbourhood neighbours. To overcome LBP's high dimensionality and low discriminative capacity in the presence of noise, many scholars have made numerous adjustments to basic LBP and produced various variations of LBP [4-8, 20, 22-25]. Image quality plays a wider role in most of the image processing applications [26]. Tan et al. [27] proposed the local ternary pattern (LTP), which stabilized the micro-patterns by employing a threshold to quantize the sign differences between the center pixel and its flanking circular pixels into three values. The LTP [27] is unaffected by changes in picture illumination. To address the noise problem, the local directional pattern (LDP) [28], is proposed and widely used in texture classification and face recognition algorithms. The LDP generates more stable edges; nonetheless, the LDP's fundamental drawback is that it requires human engagement to pick the top edge responses. This paper derived a cross diagonal LDP matrix by combining local edge responses with statistical features of texture. The proposed descriptor is a local based descriptor, with circular and elliptical structural features with an advanced version of the existing LDP.

This paper is organized as follows: the second section gives a brief idea about literature survey. The third section illustrates the proposed methodology. The section four describes the databases used for experiments and the result and analysis. The section five presents’ conclusions.

**The main contribution of this paper**

1. Derivation of a new scheme for automatically computing the top edge responses in both cross and diagonal units of isometric and anisometric edge response window. The traditional LDP requires manual intervention.
2. Derivation of ternary edge response patterns instead of binary patterns.
3. Derivation of cross and diagonal edge response ternary matrices to capture prominent features more efficiently on isometric and anisometric structures.
4. Derivation of GLCM features on the derived feature vector matrix by reducing the complexity.

**3. PROBLEM FORMULATION**

The local-based approaches on a 3x3 window have become popular for the last few
decades, especially on various classification applications. The facial skin-related applications have predominantly used local neighborhood approaches for face recognition, age classification [16, 22, 29-31], facial expression recognition, etc. Many variants of LBP, edge operators, morphological operators, etc. are derived on the 3x3 windows only.

Researchers have recently enhanced texture analysis performance by deriving feature descriptors using elliptical shapes [34, 42] rather than traditional circular, square, and rectangular shapes. The methods based on 3x3 circular patches (Fig. 1(a)) derive features only on circular windows, i.e., representing local characteristics with isotropic nature. The applications related to facial skin require the computation of elliptical or anisotropic features, since the lips, eyes, and other prominent facial skin features mostly represent elliptical shapes. In the literature the methods based on elliptical neighborhoods were not given attention, basically due to high complexity involved: to capture complete elliptical features one has to derive features on both horizontal and vertical features, which requires a third-order neighborhood with 12-pixels (fig. 1(b) and 1(c)), i.e., the complexity is doubled when compared with LBP. Recently circular and elliptical-LBP (CE-LBP) [19, 20] is proposed and proved successful in texture categorization. The advantage of CE-LBP is it captures both isotropic and complete anisotropic (circular and elliptical) features with 8-pixels by using an average function. Recently we have developed an age classification method based on CE-LBP [34] and it has shown very low error rate. The elliptical structures are used in various applications [35, 36].

The texture of a facial skin represents the various prominent features. Texture also assesses the visual structures in images and their spatial location. This research discovered no universally recognized texture definition after an in-depth examination of texture and its features. This work summarized some significant features from several texture definitions in the literature: i) The majority of the textures have the same pattern size. (ii) The texture denotes a neighborhood attribute, indicating that a pixel's intensity is heavily influenced by its surroundings. The present research assumes that the appearance of a texture is mostly determined by its illumination and the image's structure or geometry. Noise can damage the image captured by any thermal or electrical sensor, causing the actual measurement of the sensor signal to be corrupted. As a result, the sensor's signal will combine sensor signal and noise. And such images under the normal descriptors results a poor performance. Therefore, there is a need to have a local descriptor that overcomes this noise effect.

Fig. 1. Representation Of Circular And Elliptical
The major drawback of LBP is a small noise may change its value drastically. Edge detection is one of the fundamental and crucial steps for feature detection and used significantly in image processing applications. Edge-based frameworks are used for face expression recognition in the literature. Gradient magnitudes can deduce emotion-related properties of facial skin more prominently than intensity-based approaches, which is the fundamental reason for their effectiveness in face related application. The research found that histogram representation of edge-based approaches is vulnerable to noise, can produce unstable patterns, and have a negative impact on the results, particularly in the smooth regions. Edge responses exhibits better results than intensity changes. Minor noise may skew intensity values. The LDP overcomes these issues by employing a Kirsch edge response mask to generate eight edge replies per sample point [Fig. 2].

As a result, the LDP framework differs significantly from the LBP variants, which just computes the sign function between every sampling point and the center pixel. The edge responses are more resistant to noise, and changes in illumination. Furthermore, they capture more prominent information.

Fig. 2. Kirsch Edge Response Masks Are Available In Eight Different Orientations.

Fig. 3. An Image Patch Convolution With Eight Kirsch Masks Yielded The Edge Responses From 3*3
Initially, the LDP computes eight directional edge response values: $|DE_1|, |DE_2|, \ldots, |DE_7|, |DE_8|$ as given in Fig 3. The LDP creates a binary pattern based on the edge responses derived in eight directions. The basic motif behind LDP is that a direction with high edge response values contains the most structural texture information. The binary pattern is created by assigning a binary value 1 to the top n-directional edge response (DE) values out of an array of 8-DE values in the LDP. The remaining DE values (8-n) are set to zero. The LDP presumes that the top n-DE’s only include primary data. LDP’s bottleneck is determining how to select top DE’s automatically. LDP generates a binary pattern with 8-bits in the form of edge responses and it is multiplied by its corresponding $2^P$ weights, to generate a LDP code.

<table>
<thead>
<tr>
<th>$g_4$</th>
<th>$g_3$</th>
<th>$g_2$</th>
<th>80</th>
<th>67</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>$g_5$</td>
<td>$g_6$</td>
<td>$g_7$</td>
<td>46</td>
<td>56</td>
<td>57</td>
</tr>
<tr>
<td>$g_8$</td>
<td></td>
<td></td>
<td>21</td>
<td>73</td>
<td>34</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Edge responses</th>
<th>Absolute of edge responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>359 127 -57</td>
<td>359 127 57</td>
</tr>
<tr>
<td>-9 0 -321</td>
<td>9 0 321</td>
</tr>
<tr>
<td>-65 -161 127</td>
<td>65 161 127</td>
</tr>
</tbody>
</table>

**Fig. 4. The frame work of LDP with n=2.**

Figure 4 shows the framework for LDP for n=2. The LDP solves the noise-related issues that outbreaks in LBP and other intensity related operators. The LBP varies in the presence of noise, whereas the LDP maintains a consistent pattern (Fig. 5). As illustrated in Fig. 5(a), the red color pixel intensity swings due to noise. There is no change in the code generated by LDP thus it is more robust to noise (Fig. 5(b)).

The LDP is robust to noise; this present research noted the following shortcomings of LDP:

1. The code generated by LDP on a 3x3 window with 8 bits will have n number of 1’s, if one chooses top n edge responses.
2. The LDP code varies depending on the amount of top DE’s taken into account. This is seen as one of LDP’s downsides.
3. The fundamental disadvantage of LDP is that the top DE’s, i.e. the n value, must be chosen manually or randomly.
4. The accuracy of LDP-mainly dependent on the number of top DE’s. Furthermore, there is no one best method to pick randomly the best n-DE’s to improve overall performance.
5. The LDP only generates binary patterns.
6. The LDP generates a code that ranges from 0 to 255 on 3x3 window, which is similar to LBP and difficult to integrate with statistical features or any other method.
This paper proposed two cross-diagonal Local directional ternary (CD-LDT) matrices (CD-LDTM) after examining the above: one on isotropic and the other on anisometric local structures named as ICD-LDTM and ACD-LDTM respectively. The proposed descriptors consider the directional edge responses (DE) as absolute values. The DE window is divided into cross and diagonal DE windows, each of 4 DE’s. This paper automatically computes two thresholds $t_d$ and $t_c$ on diagonal and cross DE respectively (Equation 1 and 2) to represent significant DE’s. These thresholds automatically derive a ternary pattern, instead of binary patterns as basic LDP (Equation 3 and 4). Based on these ternary patterns, C-LDT and D-LDT are computed (Equation 3 and 4).

\[ t_c = \frac{\left(\sum_{i=2}^{8} DE_i\right)}{4} \quad \text{Where } i = 1,3,5,7 \]  

\[ t_d = \frac{\left(\sum_{i=1}^{7} DE_i\right)}{4} \quad \text{Where } i = 2,4,6,8 \]  

where $DE_i$ represents the $i^{th}$ directional edge response value.

The C-LDT derives the ternary pattern and C-LDT unit (C-LDT$_u$) computes a unique code

\[ C \times LDT_u = \sum_{j=1}^{l} 3^{j-1} \times S(x) \]  

\[ S(x) = \begin{cases} 
0 & \text{if } DE_i < t_c \\
1 & \text{if } DE_i = t_c \\
2 & \text{if } DE_i > t_c 
\end{cases} \quad \text{where } i = 1,3,5,7 
\]

The computation of D-LDT unit (D-LDT$_u$) is given below:

\[ D \times LDT_u = \sum_{j=1}^{l} 3^{j-1} \times S(x) \]  

\[ S(x) = \begin{cases} 
0 & \text{if } DE_i < t_d \\
1 & \text{if } DE_i = t_d \\
2 & \text{if } DE_i > t_d 
\end{cases} \quad \text{where } i = 2,4,6,8 
\]

The C-LDT$_u$ and D-LDT$_u$ ranges from 0 to 80.

This research constructed a matrix called isometric cross diagonal local direction Ternary matrix (ICD-LDTM) and anisometric cross diagonal local direction Ternary matrix (ACD-LDTM). The ICD-LDTM and ACD-LDTM are 2-D matrices of size of 81 x 81. The row of ICD-LDTM holds of C-LDT$_u$, and the columns holds D-LDT$_u$. The ICD-LDTM and ACD-LDTM derives spatial information of local edge responses on isometric and anisometric structures respectively in a precise manner. The ICD-LDTM and ACD-LDTM holds the relative frequencies of C-LDT$_u$ and D-LDT$_u$ of isometric and anisometric structures respectively. This research derived GLCM features on ICD-LDTM and thus integrated the edge response patterns with texture’s statistical features. The same process is repeated for anisometric structure in the present paper, i.e., for the derivation of ACD-LDTM. The frame work of ACD-LDTM is shown in figure 6. The quantization process is derived as given in our earlier paper [34].
Algorithm for the derivation of feature vector on the proposed descriptors: ICD-LDTM and ACD-LDTM

Algorithm begins

1. Generate the proposed ICD-LDTM and ACD-LDTM. The dimension of the matrices is 81 x 81. Initialize them to zero.
2. Consider the 3 x 3 microgrids, for isometric structures.
3. Convert the third-order neighborhood in to a 3x3 grid that holds the complete anisometric features using average function.
4. Do the following for both proposed isometric and anisometric grids.
5. Compute the directional edge responses (DE) separately using kirsch masks.
6. Compute the absolute DE’s for both structures.
7. Divide the DE values of 3x3 grid in to cross and diagonal DE’s.
8. Compute cross and diagonal thresholds $t_c$ and $t_d$.
9. Compute cross and diagonal DE-ternary pattern using $t_c$ and $t_d$.
10. Compute C-LDT, and D-LDT, on both the structure.
11. Compute ICD-LDTM(C-LDT, D-LDT) = ICD-LDTM ( C-LDT, D-LDT ) + 1
12. Compute ACD-LDTM (C-LDT, D-LDT) = ACD-LDTM ( C-LDT, D-LDT ) + 1
13. Move to the next 3 x 3 window in an overlapped manner and repeat the steps from 2 to 12. Continue this process on entire image.

14. Compute GLCM [37] features on ICD-LDTM and ACD-LDTM with different angles of rotation of 0, 45, 90, and 135 degrees to construct a feature vector of the image.

End of the algorithm

4. RESULT AND DISCUSSIONS

The suggested frameworks ICD-LDTM and ACD-LDTM were tested on the OUI-Audience Face collection, MORPH facial database, and IMDB-Wiki Dataset in this research. A brief explanation of these databases, were given in our earlier paper [34]. The sample images are shown below (Fig 7 to 9). This study derived 24-GLCM features with various rotations on proposed ICD-LDTM and ACD-LDTM. The proposed models divided people into four age groups: child (1-15 years), young (15-30 years), middle (30-60 years), and senior (over 60 years) (above 60 years). For classification, this paper used machine learning classifiers: linear regression, SVM, Decision Tree, Naive Bayes, and K-nearest Neighborhood. The Log-Loss, Area Under Curve, and F1-Score functions were used to calculate age classification accuracies in this study and displayed in the form of graphs from figures 10 to 15.

Fig 7. Sample images of OUI-Audience dataset.
The graphs 10 to 15 shows the age classification results of the proposed ICD-LDTM and ACD-LDTM on the three selected facial databases using five machine learning classifiers. The ACD-LDTM descriptor achieved high age classification rate of approximately 2 to 3% when compared to the other proposed descriptor: ICD-LDTM. The reasons for this are 1) the prominent features of facial skin mostly represent elliptical shapes and the ACD-LDTM derived edge responses and texture features on complete anisotropic structure. 2) the most desirable advantage of ACD-LDTM is it automatically represents the isometric structure also. 3) the derivation of cross and diagonal edge response thresholds simplified the overall complexity and further derived ternary patterns. The methods proposed in this paper, exhibited high age classification rate on decision tree classifier. The SVM and Linear regression stood almost second in age classification rate. The proposed ICD-LDTM achieved approximately 93%, 95% and 94%. The other proposed descriptor ACD-LDTM derived 96%, 96% and 95% average classification accuracy on OUI-Audience Face collection, MORPH facial database, and IMDB-Wiki Datasets.
Fig 8. Sample images of MORPH Album 2 dataset.

Fig 9. Imdb-Wiki Dataset.
Fig. 10: The Graph Exhibiting The Classification Accuracy On OUI-Face Data Base: ICD-LDTM.

Fig. 11: The Graph Exhibiting The Classification Accuracy On Morph-2-Face Data Base: ICD-LDTM.
**Fig. 12** The Graph Exhibiting The Classification Accuracy On OUI-Face Data Base: IMDB-Wiki: ICD-LDTM

**Fig. 13** The Graph Exhibiting The Classification Accuracy On OUI-Face Data Base: ACD-LDTM.
Fig. 14: The Graph Exhibiting The Classification Accuracy On Morph-2-Face Data Base: ACD-LDTM.

Fig. 15 The Graph Exhibiting The Classification Accuracy On OUI-Face Data Base: IMDB-Wiki: ACD-LDTM

Table 7: Classification Accuracies On Age Classification: Existing Local-Based Approaches Versus The Proposed Approaches

<table>
<thead>
<tr>
<th>S.No</th>
<th>Name of the method</th>
<th>Age Classification accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CE-LBP [20]</td>
<td>82.3</td>
</tr>
<tr>
<td>2</td>
<td>ELBP [36]</td>
<td>86.5</td>
</tr>
<tr>
<td>3</td>
<td>E. Eidinger, SerestinaViriri [38]</td>
<td>93.4</td>
</tr>
<tr>
<td>4</td>
<td>G. Levi and T. Hassner [39]</td>
<td>86.8</td>
</tr>
<tr>
<td>5</td>
<td>K. Zhang, C. Gao, L Guo et al [40]</td>
<td>93.2</td>
</tr>
<tr>
<td>6</td>
<td>Proposed ICD-LDTM</td>
<td>95</td>
</tr>
<tr>
<td>7</td>
<td>Proposed ACD-LDTM</td>
<td>98</td>
</tr>
</tbody>
</table>
This paper on each of the proposed methods computed the average classification rate by considering classification accuracies on all four age categories using a decision tree classifier and compared this with the state of art existing methods. From the experimental analysis (fig 16), this paper outlined the reasons for high age classification rate of the proposed descriptors, and given follows:

i) The computation of edge features in the form of edge directions, instead of intensity levels; ii) The computation of two different thresholds one for cross and the other for diagonal edge responses to automatically select active edge responses on cross and diagonal edge response matrices; iii) The use of a ternary pattern rather than a binary pattern to generate salient spatial information, as in the case of LBP and LDP; iv) The derivation of cross diagonal matrices to reduce the complexity and to integrate with statistical texture features; v) The derivation of cross and diagonal edge response thresholds simplified the overall complexity and further derived ternary patterns. vi) The ICD-LDTM and ACD-LDTM are not the same as the LTP (local ternary pattern) and other cross-diagonal representation matrices. The LTP computed ternary patterns using intensity comparisons, but the ternary patterns in this research are based on active edge responses; vii) Edge features, particularly directional edge response values derived in the present paper, best portray the various shapes and textures; viii) Object contours are not represented correctly by the most of the earlier methods. In areas of high contrast, the object outlines carry more discriminative information. The suggested descriptors hold the significant contour information; ix) this new coding strategy by integrating with GLCM features improved the overall age classification rate.

5. CONCLUSIONS

This paper proposes two new schemes for effective age classification based on automatic edge responses derived on cross and diagonal grid of a 3x3 window, holding isometric and anisometric properties. Further this paper derived ternary edge response patterns. This work might have not progressed well because of the complexity issues in dealing with the ternary edge response features and the complexities involved in the representation of complete anisometric structures. The main contribution of this paper is to remove completely manual or random selection of edge responses. This research addressed this by dividing the window into two separate edge response windows and separately computed the edge response thresholds to automatically determine the high edge responses without any manual intervention. This research combined the vertical and horizontal anisometric frames in to one 3x3 window and thus reduced lot of complexity involved. The derived 3x3 window holds the complete set of anisometric features. The edge response windows are divided in to cross and diagonal edge responses, facilitated for deriving a new coding scheme. The proposed two frameworks ICD-LDTM and ACD-LDTM computed the relative frequency of occurrences of cross and diagonal ternary edge.
responses isometric and anisometric frames, and further integrated with GLCM features, thus derived prominent edge, texture, and statistical features of the facial skin. The proposed derivation of ACD-LDTM automatically holds the isometric and complete anisometric features and this is the basic reason for its high classification rate than the other proposed ICD-LDTM. There is no necessity to compute ICD-LDTM, whenever the ACD-LDTM is computed. The proposed descriptors attained high age classification rate on decision tree classifier, out of the five classifiers used for classification purposes.

REFERENCES


[19] Subba Reddy, K., Vijaya Kumar, V., Siva Kumar, A.P., “Classification of textures using a new descriptor circular and elliptical-LBP


[34] Akarapu Swarna, Venkata Krishna, Supreethi K.P; Age Classification Based on Relative-Features Derived on Circular and Elliptical Neighborhoods using Machine Learning Classifiers ; Design Engineering ISSN: 0011-9342 | Year 2021 Issue: 9


