Binary Pattern Analysis for 3D Facial Action Unit Detection

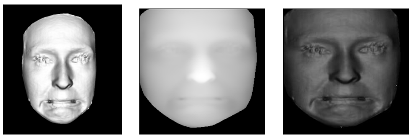
Georgia Sandbach[1]
gls09@imperial.ac.uk
Stefanos Zafeiriou[1]
s.zafeiriou@imperial.ac.uk
Maja Pantic[1,2]
m.pantic@imperial.ac.uk

[1] Department of Computing
Imperial College London
London, UK

[2] EEMCS
University of Twente
Enschede, Netherlands

Recognition of facial expressions is a challenging problem, as the face is capable of complex motions, and the range of possible expressions is extremely wide. For this reason, detection of facial action units (AUs) from the Facial Action Coding System, a comprehensive system for coding facial muscle movements, has become a widely studied area of research. The use of 3D facial geometry data and extracted 3D features for expression recognition has so far not been heavily studied. Images and videos of this kind allow a greater amount of information to be captured (2D and 3D), including out-of-plane movement which 2D cannot record, whilst also removing the problems of illumination and pose inherent to 2D data. For this reason some work has begun to employ 3D facial geometry data for facial expression recognition or facial AU detection.

We tackle this problem with the introduction of a variety of new binary pattern features that are all based on the traditional Local Binary Pattern (LBP) [6] or Local Phase Quantiser (LPQ) [5] features. In order to do this, we employ two 2D representations of the 3D facial geometry information. Firstly we utilise the depth map, that has been widely used in 3D facial analysis, and secondly we define the Azimuthal Projection Distance Image (APDI), which captures the comparative directional information of the normals in the mesh as a 2D representation. The Azimuthal Equidistant Projection (AEP) is able to project normals onto positions in a Euclidean 2D plane. For our purposes we alter the projection to create the APDI, which allows direct comparison of the projection coordinates of neighbouring points.



(a) (b) (c)

Figure 1: 2D representations of the facial mesh for subject AU20 (a) Original facial mesh (b) Depth map (b) Azimuthal Projection Distance Image

We have a regular grid of normals, defined as $\mathbf{n}(i,j) = (u_{i,j}, v_{i,j}, w_{i,j})$, to be projected relative to a set of mean normals $\hat{\mathbf{n}}(i,j)$ calculated at each point. We set the elevation and azimuth of the mean normals, $\hat{\theta}$ and $\hat{\phi}$, to be $\frac{\pi}{2}$ and 0 respectively at every point. This then allows calculation of the distances of the normals as compared to the mean $\hat{\mathbf{n}} = (1,0,0)$ which is chosen as a reference to create an image suitable for further analysis. This assumption allows the AEP projection to be simplified to:

$$x_{i,j} = k' cos\theta(i,j) sin\phi(i,j) \qquad y_{i,j} = k' cos\theta(i,j) cos\phi(i,j) \tag{1}$$

for a point $\mathbf{p}(i,j) = (x_{i,j}, y_{i,j})$, where θ is the elevation angle, and ϕ is the azimuthal angle, of the normal at this point. The above formulation allows comparison between normal distances in Euclidean space, and this simplification also reduces the complexity of the feature extraction process. In order to employ this in the binary pattern framework, the coordinates are used to find an absolute distance from the origin $d_{i,j} = \sqrt{x_{i,j}^2 + y_{i,j}^2}$, and these values form the APDI for the facial mesh. Examples of the depth map and APDI, calculated for the facial mesh seen in Fig. 1(a), are shown in Figs. 1(b) and 1(c) respectively.

Each representation is then exploited for use with binary pattern algorithms in order to form feature types suitable for robust AU detection. Firstly, the traditional Local Binary Pattern (LBP) algorithm, which assigns a binary pattern to each point in an image by thresholding the neighbouring points on the central value, was applied directly to each representation. This forms the previously proposed 3DLBP [7], for use as a baseline test, and the new Local Azimuthal Binary Pattern (LABP) respectively. Next we utilise other methods that have been employed for

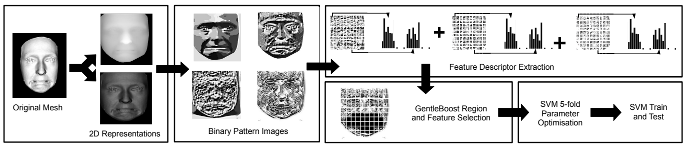


Figure 2: An overview of our proposed system.

2D feature extraction: LPQs, Local Gabor Binary Patterns (LGBPs) [4], and Histogram of Monogenic Binary Patterns (HMBPs) [3]. In the latter two cases, we extend the methods to apply them to the magnitude, phase, and, in the case of the Monogenic signal, orientation. In total this formed seven new features: (1) Local Azimuthal Binary Patterns (LABPs) (2) Local Depth Phase Quantisers (LDPQs) (3) Local Azimuthal Phase Quantisers (LAPQs) (4) Local Depth Gabor Binary Patterns (LDGBPs) (5) Local Azimuthal Gabor Binary Patterns (LAGBPs) (6) Local Depth Monogenic Binary Patterns (LDMBPs) (7) Local Azimuthal Monogenic Binary Patterns (LAMBPs). The performance of these new features is assessed as compared to the original 3DLBPs.

Feature vectors are created for each of the above descriptors through the use of histograms. First, the *x*-*y* plane of the mesh is divided into 10×10 equally-sized square blocks, and for each of these a histogram is formed from the calculated binary numbers. These histograms are then concatenated into one large feature vector. Feature selection is performed in order to reduce the dimensionality of the feature vectors before classification. The GentleBoost algorithm was used for this purpose, with two stages to the feature selection: first selection of regions, and then particular features. Support Vector Machines (SVMs) were then trained for detection of each AU, with parameter optimisation carried out using 5-fold cross-validation, and these were used for testing of all sequences. An overview of our system can be seen in Fig. 2.

Experimental testing was conducted in two ways: 10-fold cross- validation on the Bosphorus database [2], and cross-database testing with training on this database, and testing carried out on the D3DFACS database [1]. The results achieved show a definite improvement with all of the new features over the traditional 3DLBP, with a maximum cross-validation ROC AuC of 97.2 when using LDGBPs. This improvement was also seen with the depth features on the cross-database testing, though the Azimuth results in this test suggested that these features are less robust when there are large variations in smoothness of the mesh between training and testing data.

[1] D. Cosker et al. A FACS valid 3D dynamic action unit database. In *ICCV 2011*, pages 2296–2303. IEEE, 2011.

[2] A. Savran et al. Bosphorus database for 3D face analysis. *Biometrics and Identity Management*, pages 47–56, 2008.

[3] M. Yang et al. Monogenic Binary Pattern (MBP): A Novel Feature Extraction Model. In *ICPR 2010*, pages 2680–2683. Ieee, 2010.

[4] W. Zhang et al. Local Gabor binary pattern histogram sequence (LGBPHS). In *ICCV 2005*, volume 1, pages 786–791. IEEE, 2005.

[5] V. Ojansivu and J. Heikkilä. Blur insensitive texture classification using local phase quantization. *ISP*, pages 236–243, 2008.

[6] T. Ojala et al. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE PAMI*, pages 971–987, 2002.

[7] Y. Huang et al. Combining statistics of geometrical and correlative features for 3D face recognition. In *BMVC 2006*, pages 879–888. Citeseer, 2006.