

# Riesz-based Volume Local Binary Pattern and A Novel Group Expression Model for Group Happiness Intensity Analysis

Xiaohua Huang<sup>1</sup>  
 huang.xiaohua@ee.oulu.fi  
 Abhinav Dhall<sup>2,3</sup>  
 abhinav.dhall@canberra.edu.au  
 Guoying Zhao<sup>1</sup>  
 gyzhao@ee.oulu.fi  
 Roland Goecke<sup>2,3</sup>  
 roland.goecke@ieee.org  
 Matti Pietikäinen<sup>1</sup>  
 mkp@ee.oulu.fi

<sup>1</sup> Center for Machine Vision Research  
 Department of Computer Science and Engineering  
 University of Oulu, Finland  
<sup>2</sup> Human-Centred Technology Research Centre  
 University of Canberra, Australia  
<sup>3</sup> IHCC Group  
 Research School of Computer Science  
 Australian National University, Australia

Automatic emotion analysis and understanding has received much attention over the years in affective computing. This paper aims to infer the emotional intensity of a group of people. For group emotional intensity analysis [1], feature extraction and group expression model (GEM) are two critical issues.

**Feature extraction:** For 2D image, the 1st-order Riesz transform enables the rotationally invariant analysis of 1D structure, while the 2nd-order Riesz transform can characterize an i2D image structure [4]. Therefore, we employ the 1st-order and 2nd-order Riesz transforms with log-Gabor filters at three scales and four orientations to describe the structure of a facial expression image.

Its Riesz face images are resembled to form the Riesz volume. We furthermore employ the LBP operator on  $XY$ ,  $XZ$  and  $YZ$  of volume-based Riesz face respectively. The histograms  $H_{XY}$ ,  $H_{XZ}$  and  $H_{YZ}$  are concatenated into one feature vector  $H$ . The procedure is shown in Fig. 1. Five histograms  $H_x, H_y, H_{xx}, H_{xy}, H_{yy}$  are concatenated into one feature vector  $F$  for representing the face. Based on a low-dimensional space  $U$ , the low-dimensional features can be formulated as  $\tilde{F} = U^T F$ .

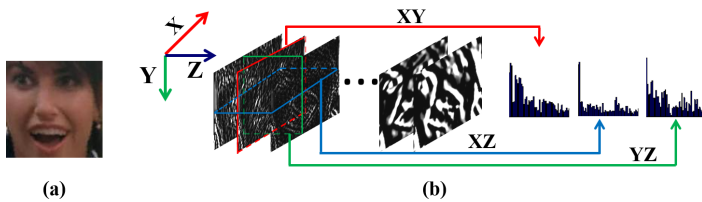


Figure 1: Riesz-based volume local binary patterns for  $R_x$ : (a) A face image; (b) The 1st-order Riesz face  $R_x$  and its feature extraction.

**Group expression model:** GEM explores the relationship between faces and intensity label in a group image. We formulate a novel GEM based on Continuous Conditional Random Fields (GEM<sub>CCRF</sub>) to combine global and local attributes.

For the global attribute, the minimal spanning tree algorithm [3] is employed to obtain a fully connected graph  $\mathcal{G} = (V, E)$ , which can provide the location and minimally connected neighbours of a face. An example is shown in Fig. 2. Based on the graph, the global attribute can be expressed by relative size  $S_i$  and relative distance  $\delta_i$ . For the local attribute, it contains the local features  $\tilde{F}$  for faces.

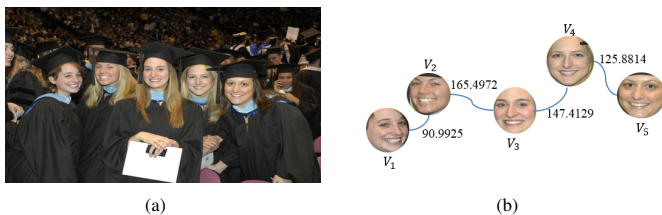


Figure 2: Illustration of obtaining minimally connected neighbours for a face: (a) Original group image, (b) Min-span tree depicting connection between faces.

Global and local attributes are concatenated into  $\mathcal{F} = [\tilde{F}, S, \delta]$ , providing sufficient and useful information for CCRF. We suppose that a group image contains  $n$  faces, which is represented by  $\mathcal{F}_i, i = 1, \dots, n$ . We transform  $\mathcal{F}_i$  into  $t_i$  by using Support Vector Regression, which represents the relevance factor of one subject for CCRF. Therefore, CCRF model for a group image is a conditional probability distribution with the probabil-

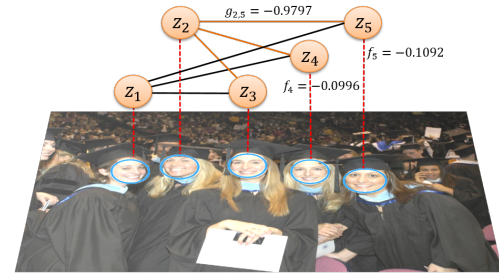


Figure 3: GEM<sub>CCRF</sub>: The blue circle on a facial image represents the extracted content,  $z_i$  is a happiness intensity label, an edge (a solid line) between  $z_i$  and  $z_j$  means the dependency between intensities of two faces, an edge (a dash line) represents the dependency of an intensity label on its content.



Figure 4: Examples of inferring two group images on HAPPEI database. For each image the texts at the bottom of an image indicate the estimated happiness intensity results and the ground truth, respectively

ity density function as  $P(z | \mathbf{T}) = \frac{\exp(\sum_{i=1}^n \sum_{k=1}^m \mu_k f_k(z_i, \mathbf{T}_{i,k}) + \sum_{i,j} v g(z_i, z_j, \mathbf{T}))}{\int \exp(\sum_{i=1}^n \sum_{k=1}^m \mu_k f_k(z_i, \mathbf{T}_{i,k}) + \sum_{i,j} v g(z_i, z_j, \mathbf{T})) dz}$ , where  $\mathbf{T}$  is the set of input feature vector,  $\mathbf{Z}$  is the intensity label of faces in a group image,  $f(\cdot)$  and  $g(\cdot)$  are vertex and edge features, respectively. Following the work of [2], we pick  $\mu$  and  $v$  that maximise the conditional log-likelihood  $(\mu^*, v^*) = \arg \max_{\mu, v} \sum_{q=1}^M \log P(\mathbf{z}^{(q)} | \mathbf{T}^{(q)})$ , where  $M$  expresses the number of group images.

Given a group image, containing  $n$  faces, the intensity can be estimated by using  $\mathbf{z} = \frac{1}{n} \sum_{i=1}^n \arg \max_{z_i} P(z_i | \mathbf{T}_{test})$ . An example of the CCRF model in group happiness intensity analysis is depicted in Fig. 3.

Implementation of this framework by group-level happiness intensity estimation, as shown in Fig. 4, is described in the paper. Our conclusion is that GEM<sub>CCRF</sub> helps to predict the group mood better. Moreover, the feature indeed affects the performance of recognition for GEM models.

[1] A. Dhall, R. Goecke, and T. Gedeon. Automatic group happiness intensity analysis. *IEEE Transaction on Affective Computing*, 6(1): 13–26, 2015.  
 [2] V. Imbrsaitė, T. Baltrušaitis, and P. Robinson. Emotion tracking in music using continuous conditional random fields and relative feature representation. In *IEEE International Conference on Multimedia and Expo Workshops*, pages 1–6, 2013.  
 [3] R. Prim. Shortest connection networks and some generalizations. *Bell System Technical Journal*, 36(6):1389–1401, 1957.  
 [4] L. Zhang and H. Li. Encoding local image patterns using Riesz transform: With applications to palmpring and finger-knuckle-print recognition. *Image and Vision Computing*, 30(12):1043–1051, 2012.