**SPECIAL ISSUE PAPER**

# Detecting personality and emotion traits in crowds from video sequences

Rodolfo Migon Favaretto[1] · Paulo Knob[1] · Soraia Raupp Musse[1] · Felipe Vilanova[2] · Ângelo Brandelli Costa[2]

**Abstract**

This paper presents a methodology to detect personality and basic emotion characteristics of crowds in video sequences. Firstly, individuals are detected and tracked, and then groups are recognized and characterized. Such information is then mapped to OCEAN dimensions, used to find out personality and emotion in videos, based on OCC emotion models. Although it is a clear challenge to validate our results with real life experiments, we evaluate our method with the available literature information regarding OCEAN values of different Countries and also emergent Personal distance among people. Hence, such analysis refer to cultural differences of each country too. Our results indicate that this model generates coherent information when compared to data provided in available literature, as shown in qualitative and quantitative results.

**Keywords** Computer vision · Crowd features · Big Five model · Cultural dimensions · Crowd emotion

## 1 Introduction

Crowd analysis is a phenomenon of great interest in current applications. Surveillance, entertainment, and social sciences are examples of fields that can benefit from the development of this area of study. Literature presents different applications of crowd analysis, like counting people in crowds [5,8], group and crowd movement and formation [28,38,41,45] and detection of social groups in crowds [9,22,39,40]. Normally, these approaches are based on personal tracking or optical flow algorithms, and handle with features like walking speed, directions and distances over time. Specifically on this subject, one study investigated cultural difference in videos from different countries: Chattaraj et al. [10] suggested that cultural and population differences could produce deviations in speed, density and flow of the crowd.

In this paper, we propose to detect personality aspects based on the Big Five personality model, also referenced as (OCEAN) [14], using individuals behaviors automatically detected in video sequences. For this, we used the NEO PI-R [14] which is the standard questionnaire measure of the five-factor model. Also, we use such psychological traits to identify some primordial emotions in the crowd. Using a similar mapping as proposed by Saifi et al. [37], we were able to identify the level of some emotions for each individual, like happiness or fear, according its OCEAN level.

This paper is organized as follows: Sect. 2 presents related works on personality and crowd emotion detection. Section 3 describes our model to detect personality traits for individuals and groups. Moreover, we use this psychological traits to identify different levels of emotions in the crowd, like if an agent is happy or angry, in Sect. 3.3. Experimental results are addressed in Sect. 4, where we deal with the challenge to evaluate our approach with real life. Conclusions and future work are presented in Sect. 5.

## 2 Related work

This section discusses some topics concerned with personality and emotion detection in crowds domain.

✉ Rodolfo Migon Favaretto
rodolfo.favaretto@acad.pucrs.br

[1] VHLab, Graduate Program in Computer Science, PUCRS, Av Ipiranga, 6681, Porto Alegre, RS, Brazil

[2] Graduate Program in Psychology, PUCRS, Av Ipiranga, 6681, Porto Alegre, RS, Brazil

Personality may be labeled as deep psychological individual-level trait [7]. Trait is an inference made after observed behaviors that seeks to explain its regularity [27]. Raymond Cattel is commonly referred as the one who developed the methodology which permitted the objective grouping of hundreds of trait descriptors in a set of higher level factors [17]. Cattell [6] developed a taxonomy of individual differences that consisted of 16 primary factors and 8 second-order factors. Nevertheless, attempts to replicate his work were unsuccessful [23] and researchers agreed that only the five-factor model matched his data, originating the Big Five personality model.

Nowadays, researchers agree that there are five robust orthogonal traits which effectively matched personality attributes [17], known as the Big Five: openness to experience (the active seeking and appreciation of new experiences); conscientiousness (degree of organization, persistence, control and motivation in goal directed behavior); extraversion (quantity and intensity of energy directed outwards in the social world); agreeableness (the kinds of interaction an individual prefers from compassion to tough mindedness); neuroticism (how much prone to psychological distress the individual is) [30].

The NEO PI-R [13] is one of the most used instrument based on the Big Five personality theory. It assesses the normal adult personality and is internationally recognized as a gold standard for personality assessment. One of its advantages is that it further specifies six facets within each personality trait and have data from several countries which easily allows cross-cultural comparisons [31,32]. Although the empirical evidence matching individual-level traits and crowd behavior is not strong (one of the few examples is [2]), the Big Five personality model is widely used to model computational crowd simulation [18,25,29]. In general, in such methods, OCEAN model allows to simulate a crowd with individual level parameters based on the expected behaviors of the agents.

Several models have been developed to explain and quantify basic emotions in humans and other animals. One of the most cited is the model proposed by Paul Ekman [19] which considers the existence of 6 universal emotions based on cross-cultural facial expressions (anger, disgust, fear, happiness, sadness and surprise). Other approaches such as affective neuroscience postulate, from an evolutionary perspective, consider other groups of emotions such as fear, rage/anger, and sadness/panic [33].

Panksepps theory claims that individual differences in primary emotional systems may represent the phylogenetically oldest parts of human personality. In order to support this idea, the author links individual differences in primary emotions and the Big Five model of personality. for example, seeking is robustly linked with openness, high play with higher extraversion, high care and low anger are associated with higher agreeableness, high scores on fear, sadness and rage with high neuroticism [15,16]. Finally, other models such as the proposed by Ortony, Clore, and Collins (commonly referred to as the OCC model) distinguish 22 emotion types manly bases on its expression. This is the model used in this paper [12].

In the work proposed by Gorbova and collaborators [24], authors present a system of automatic personality screening from video presentations in order to make a decision whether a person has to be invited to a job interview. The automatic personality screening is based on visual, audio and lexical cues from short video clips. The system is build to predict candidate scores of Big Five personality traits and to estimate a final decision, to which degree the person from video clip has to be invited to the job interview.

Baig et al. [1] focus their work on the perception of emotion due to crowd behavior. For this end, they propose an approach based on probabilistic modeling, which is trained to perceive the emotions of people in a given area. There, it uses camera sensors in order to track the motion of the individuals in the crowd. They use data mining techniques to identify different behaviors and events. Agents can experience different types of interactions during the simulation, which can arouse negative or positive emotions. Tests showed that the algorithms for emotions detection performed well with all tested scenarios and types of interaction. Yet, they comment that it is still necessary to include more complex interactions and motivations in a more complex probabilistic model in order to have a more realistic model.

Following the emotion detection in crowds subject, Rabiee et al. [35] mention that emotions can be valuable traits in the understanding of the crowd. Moreover, they say that the majority of methods proposed are just based on low-level visual features, leaving a huge semantic gap between these low-level features and high-level concept of crowd behavior. Therefore, in their work, it is proposed an attribute-based strategy to generate a crowd dataset with both annotations of abnormal crowd behavior and crowd emotion. In fact, the dataset is the main contribution of their work, which can be used as a benchmark for computer vision as well to help to understand the correlation between crowd behavior and emotion.

Saifi et al. [37] proposed a mapping between psychological traits and emotions. In their work, they aim to model emergent emotions and behaviors in a simulated crowd, based on the personality of each individual. For this, the OCEAN [14] personality model is combined with the OCC [34] emotional model in order to find the susceptibility of each of the five OCEAN personality factors to feel each OCC emotion. They use fuzzy logic to model the critical emotions when in presence of an unexpected event, which can trigger a specific behavior. In their method, it is possible
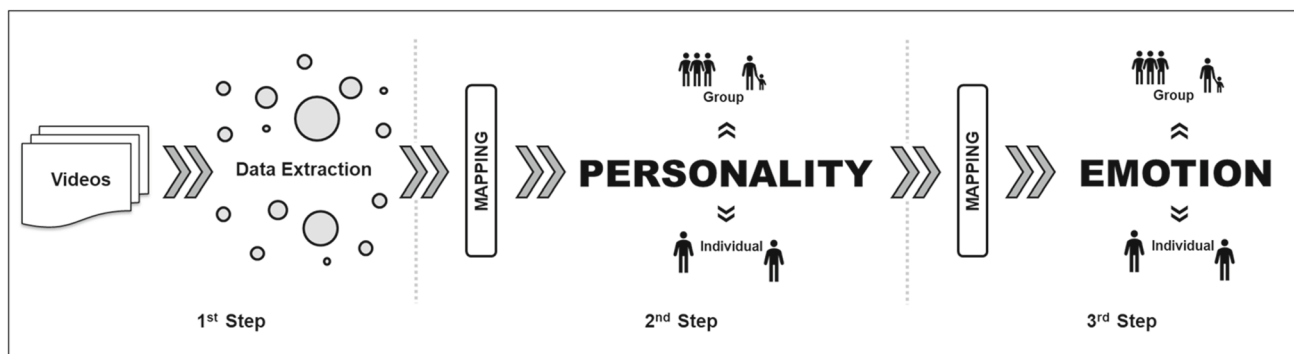
**Fig. 1** Three main steps of our approach: (1st) individual tracking and data extraction, (2nd) individual data is mapped to individual and group personality (OCEAN) traits, and (3rd) personality traits are mapped to individual and group basic emotions. These steps are detailed in Fig. 2

to observe a crowd and the effect of events, like how calm or nervous agents react to it. Moreover, the user is able to predict future situations or events, which can lead him/her to make better decisions at the right time.

Following a similar way, Zhang et al. [44] proposed a novel crowd representation named "crowd mood", which should be based on the spacing interactions of the individuals and the structural levels of motion patterns in crowds. They affirm that "Basic types of crowd motion can reflect representative emotions of the crowd.". Therefore, the main idea is to project varied types of statistical motion features and, based on this, assign emotional labels to the behavior of the crowd. The preliminary achieved results showed an interesting performance in the given tasks, offering a promising tool which represents crowd behaviors semantics and can be applied to varied crowd tasks.

In this paper, the idea is to map parameters from individual and group behaviors automatically detected from video sequences of different countries, to OCEAN dimensions. We extended a previous work [21] in order to detect groups and find their OCEAN levels as well. Then, we use OCEAN of groups and individuals to find out their emotion in the video sequence, following a mapping similar with the one proposed by Saifi et al. [37]. While their work present a model to simulate crowds with psychological traits and emotions, our work focus on identifying such traits in video sequences. Further details are going to be presented in Sect. 3. In this sense, our contribution is a model based on a set of equations that handle the individual parameters obtained from videos and map them to the methodology that compose the Big Five personality model for individuals and groups, proposing a mapping function to find out emotion values. Our method does not require any training or specific dataset since we based our prototype on the theories behind the OCEAN and OCC (emotion model).

## 3 The proposed approach

Our model presents three main steps responsible for following: video data extraction, personality and emotion analysis. These steps are illustrated in the overview of the method in Fig. 1. The first step aims to obtain the individual trajectories from observed pedestrians in real videos. Using these trajectories, we detect groups and extract data which are useful for second step, that is responsible for personality analysis of groups and individuals. Once we have concluded the second step, we have enough information to follow with the third step, which consists of emotion detection of individuals and groups according to OCEAN values. Figure 2 illustrates a more detailed flow chart of our approach and Sect. 3.1 presents further details.

### 3.1 Individuals and groups data extraction

Initially, the information about people from real videos is obtained using a tracker [3] to recover people trajectories. In order to transform image in world coordinates, we assume that the head position is on the ground place ($z = 0$), and then we used a *planar homography* to rectify a perspective image and generate world coordinates. In computer vision, *planar homography* is defined as a projective mapping from one plane to other. Then we use the homography to rectify a perspective image, in this case, to generate a "plan" view of trajectories from a "perspective" photograph. Figure 3 illustrates the mapping of the trajectories points from image coordinates to a orthogonal plan view (world coordinates). The coordinates of trajectories are used to calculate the motion parameters (speed of travel, direction) for each agent.

We compute information for each person $i$ at each timestep: (i) 2D position $\mathbf{x}_i$ (meters); (ii) speed $s_i$ (meters/frame); (iii) angular variation $\alpha_i$ (degrees) w.r.t. a reference vector $\mathbf{r} = (1, 0)$; (iv) isolation level $\varphi_i$; (v) socialization
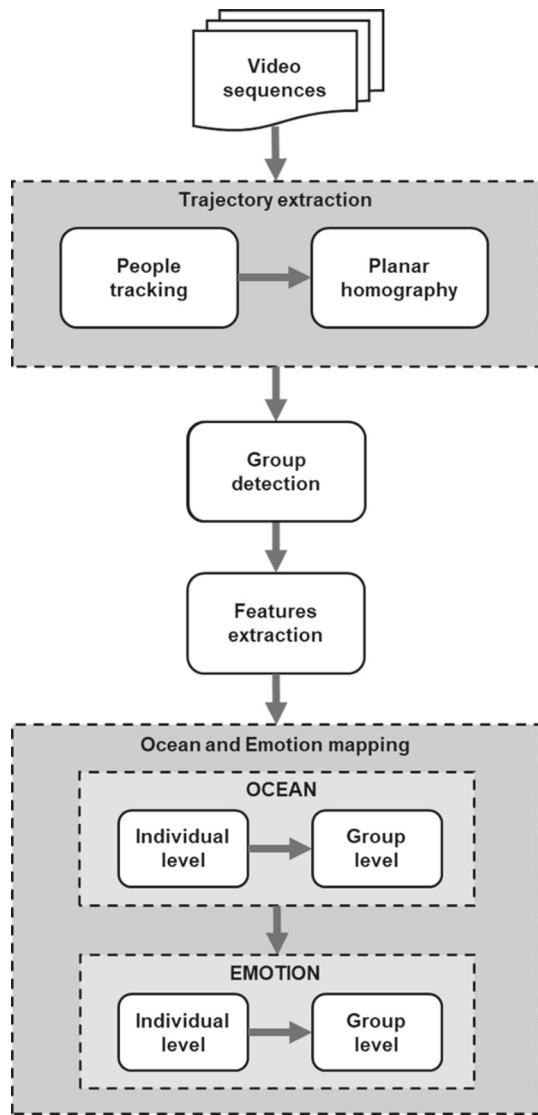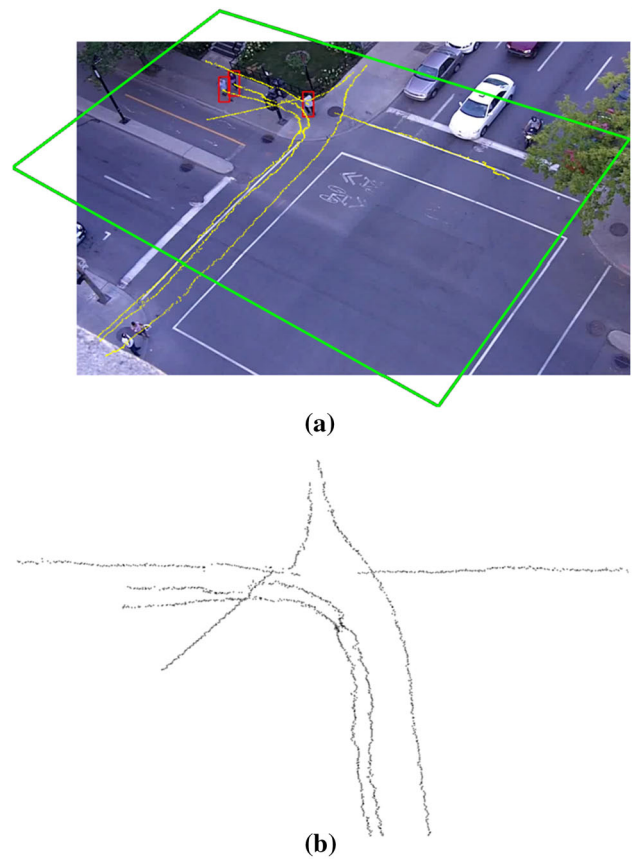
**Fig. 2** Flowchart of our approach: initially, people trajectories are extracted and passed by a homography transformation process. Based on these trajectories, groups are detected and its features are extracted. These features are the inputs for the OCEAN and EMOTION acquisition process



**Fig. 3** Homography transformation process: **a** people trajectories (yellow points in the scene) are obtained from the tracker process and rectified by homography transformation process (**b**) (color figure online)

level $\vartheta_i$; and (vi) collectivity $\phi_i$. The isolation level $\varphi$ is computed as shown in Eq. 1:

$$\varphi_i = \begin{cases} 1, & \text{if } n_{\text{social}} = 0 \\ \dfrac{\frac{1}{n_{\text{social}}} \sum_{j=0}^{n_{\text{social}}-1} d(\mathbf{x}_i, \mathbf{x}_j)}{d_{\text{Hall}}}, & \text{otherwise} \end{cases}, \quad (1)$$

where $n_{\text{social}}$ is the number of individuals in the social space[1] according to Hall's proxemics [27], $d$ is a simple function to calculate the Euclidean distance between two individuals $i$ and $j$ ($\mathbf{x}_i$ and $\mathbf{x}_j$ are, respectively, the positions of individuals

[1] Social space is related to 3.6 m [27].

$i$ and $j$) and $d_{\text{Hall}} = 3.6$ is the distance (in meters) around an individual that represents its personal social space. The socialization level $\vartheta_i$ of an individual $i$ is calculated according to Eq. 2:

$$\vartheta_i = \begin{cases} 0, & \text{if } n_{\text{social}} = 0 \\ \dfrac{n_{\text{social}}}{\rho}, & \text{otherwise} \end{cases}, \quad (2)$$

where $\rho$ is the total number of individuals in the analyzed frame. To compute the collectivity affecting individual $i$ and exerted by $n_{\text{social}}$ individuals in social space (as presented in [20]), we use Eq. 3, as follow:

$$\phi_i = \sum_{j=0}^{n-1} \gamma e^{(-\beta \varpi(i,j)^2)}, \quad (3)$$

where the collectivity between two individuals $i$ and $j$ is calculated as a decay function of $\varpi(i,j) = s(s_i, s_j).w_1 + o(\alpha_i, \alpha_j).w_2$, considering $s$ and $o$ respectively their speed and orientation differences, and $w_1$ and $w_2$ are constants that should regulate the offset in meters and radians. We have used $w_1 = 1$ and $w_2 = 1$. $\gamma = 1$ is the maximum collectivity

value when $\varpi(i, j) = 0$, and $\beta = 0.3$ is empirically defined as decay constant. Hence, $\phi_i$ is a value in the interval $[0; 1]$. Therefore, for each individual $i$ at each frame $f$ in a video sequence, we compute vector $\mathbf{V_{i,f}}$ of extracted data where $\mathbf{V_{i,f}} = [s_{i,f}, \alpha_{i,f}, \varphi_{i,f}, \vartheta_{i,f}, \phi_{i,f}]$. Yet, the average values for all frames in the video sequence are represented through the vector $\mathbf{V_i}$ of individual $i$ where $\mathbf{V_i} = [\bar{s}_i, \bar{\alpha}_i, \bar{\varphi}_i, \bar{\vartheta}_i, \bar{\phi}_i]$.

For groups detection, we use the computed parameters $s$, $o$ and $d$ which respectively state for speed, orientation variation, and distance of each pair of agents $i$ and $j$. We use the notion of distances based on the *proxemics* described by Hall [26] to define that two agents belong to the same group according with three tests empirically defined: If $(d(x_i, x_j) \leq 1.2\,\text{m})$ and $(o(\alpha_i, \alpha_j) \leq 15°)$ and $(s(s_i, s_j) < \beta)$, where $\beta = 5\%$ of higher speed. Based on this set of rules, agents are grouped in pairs. In a next step, we check which pairs have one individual in common, and merge them into larger groups. This process is performed until the group formation does not share individuals, i.e., they are disjoint. For each group $g$ detected in video sequence, we define a vector $\mathbf{G_g} = [n_g, \mathbf{I}_g, \bar{s}_g, \bar{\alpha}_g, \bar{d}_g]$ of extracted data, where $n_g$ is the number of individuals in $g$, $\mathbf{I}_g$ states for a vector containing the id of each individual in $g$ and $\bar{s}_g$, $\bar{\alpha}_g$ and $\bar{d}_g$ present the average values of speed, orientation, and distance applied by individuals in group $g$ during the video sequence.

At the end of this step, we have $\mathbf{V}$ for all individuals per frame and averaged for the video sequence and $\mathbf{G}$ for all groups averaged for the video sequence. In next section we describe how these data is used to find out the personality traits in our work.

## 3.2 Mapping crowd features in OCEAN dimensions

We used the empirically equations proposed by Favaretto et al. [21] in a previous work which goal is to map individual and group characteristics in OCEAN cultural dimensions. Basically, the method proposes to "answer" 25 items from NEO PI-R inventory for each individual in the video sequence. Although the complete version of NEO PI-R has 240 items, 25 of them were selected since they have a direct relationship with crowd behavior. Each question is mapped to an equation using only information contained in $V$ for each $i$. For example, in order to represent the item "1. Have clear goals, work to them in orderly way", we consider that the individual $i$ should have a high velocity $s$ and low angular variation $\alpha$ to have answer compatible with "Strong". So the equation for this item is $Q_1 = s_i + \frac{1}{\alpha_i}$. Table 1 shows the equations used for each considered question.

Once all questions $k$ (in the interval $[1; 25]$) have been answered for all individuals $i$, we have $\mathbf{Q^f_{i,k}}$ for each frame $f$. In addition, we computed the average values to have one vector $\mathbf{Q_{i,k}}$ per video. According to NEO PI-R definition, each of the questions $\mathbf{Q'_k}$ are associated to one of the Big

Five dimensions and some questions should invert the values, because an item score 4 (strongly agree) can represent a high or low value of a certain personality trait. So, to get the correct values, we applied a factor to the questions which score should be inverted: $\mathbf{Q^*_{i,k}} = 4 - \mathbf{Q'_{i,k}}$, as shown in next equations:

$$O_i = \frac{Q^*_{i,2}}{\varrho}, \tag{4}$$

$$C_i = \frac{Q'_{i,1}}{\varrho}, \tag{5}$$

$$E'_i = Q'_{i,3} + Q'_{i,12} + Q'_{i,14} + \sum_{q=16}^{23} Q'_{i,q}, \tag{6}$$

$$E^*_i = \sum_{q=4}^{8} Q^*_{i,q} + Q^*_{i,11} + Q^*_{i,15}, \tag{7}$$

$$E_i = \frac{(E'_i + E^*_i)}{\varrho}, \tag{8}$$

$$A_i = \frac{\sum_{q=9}^{10} Q'_{i,q}}{\varrho}, \tag{9}$$

$$N_i = \frac{Q'_{i,13} + \sum_{q=24}^{25} Q^*_{i,q}}{\varrho}, \tag{10}$$

where $\varrho$ represents the percentage of questions from the total, in each dimension (O, C, E, A and N), respectively 4%, 4%, 72%, 8% and 12%.

## 3.3 Emotion detection

In this paper we connect personality and emotion traits in order to detect data in the individuals in video sequences. As mentioned by [36]:

"A helpful analogy is to consider that personality is to emotion as climate is to weather. That is, what one expects is personality, what one observes at any particular moment is emotion." That is the reason why, in our work, we correlated emotion traits with visual behaviors that will be perceived in pedestrian motion.

Our work proposes an emotion mapping based on personality traits (i.e., OCEAN) found for each individual present in the video sequence. First, we selected four emotions from OCC [34] model: Fear, Happiness, Sadness and Anger. It is important to notice that we chose only four emotions that, in our opinion, are the most visible when relating with motion behavior, which is our case in video sequences with pedestrians. Any other from the total of 22 emotions proposed in OCC model could also be mapped.

Once we have computed the personality traits for an individual, we propose a way to map these personality traits into the four considered emotions. In addition to

**Table 1** Equations from each NEO PI-R selected item

| NEO PI-R Item | Equation |
|---|---|
| 1. Have clear goals, work to them in orderly way | $Q_1 = s_i + \frac{1}{\alpha_i}$ |
| 2. Follow same route when go somewhere | $Q_2 = \alpha_i$ |
| 3. Shy away from crowds | $Q_{3-8} = \varphi_i$ |
| 4. Dont get much pleasure chatting with people | |
| 5. Usually prefer to do things alone | |
| 6. Prefer jobs that let me work alone, unbothered | |
| 7. Wouldnt enjoy holiday in Las Vegas | |
| 8. Many think of me as somewhat cold, distant | |
| 9. Rather cooperate with others than compete | $Q_{9-10} = \phi_i$ |
| 10. Try to be courteous to everyone I meet | |
| 11. Social gatherings usually bore me | $Q_{11} = \varphi_i + \text{std}(\alpha_i)$ |
| 12. Usually seem in hurry | $Q_{12} = s_i + \alpha_i$ |
| 13. Often disgusted with people I have to deal with | $Q_{13} = \varphi_i + \frac{1}{\phi_i}$ |
| 14. Have often been leader of groups belonged to | $Q_{14} = \phi_i + \vartheta_i + \frac{1}{\alpha_i}$ |
| 15. Would rather go my own way than be a leader | $Q_{15} = \frac{1}{Q_{14}}$ |
| 16. Like to have lots of people around me | $Q_{16-21} = \vartheta_i$ |
| 17. Enjoy parties with lots of people | |
| 18. Like being part of crowd at sporting events | |
| 19. Would rather a popular beach than isolated cabin | |
| 20. Really enjoy talking to people | |
| 21. Like to be where action is | |
| 22. Feel need for other people if by myself for long | $Q_{22-25} = \vartheta_i + \phi_i$ |
| 23. Find it easy to smile, be outgoing with strangers | |
| 24. Rarely feel lonely or blue | |
| 25. Seldom feel self-conscious around people | |

vector $\mathbf{V}_{i,f}$ for individual $i$ per frame $f$, we included personality and emotion parameters as follows: $\mathbf{V}_{i,f} = [s_{i,f}, \alpha_{i,f}, \varphi_{i,f}, \vartheta_{i,f}, \phi_{i,f}, \mathbf{P}_{i,f}, \mathbf{E}_{i,f}]$. $\mathbf{P}_{i,f}$ states for values of OCEAN personality $[O_{i,f}, C_{i,f}, E_{i,f}, A_{i,f}, N_{i,f},]$, while $\mathbf{E}_{i,f}$ states for values of emotion based on OCC model: $[F_{i,f}, H_{i,f}, S_{i,f}, An_{i,f}]$. Values of $\mathbf{P}$ are in the interval $[0; 1]$ while values of $\mathbf{E}$ are from $[-3; 3]$. For groups analysis, we only considered parameters of personality and emotion in the detected groups $g$ in the video sequence. In addition to previously defined $\mathbf{G_g} = [n_g, \mathbf{I}_g, \bar{s}_g, \bar{\alpha}_g, \bar{d}_g]$, we include $\mathbf{P}_g$ and $\mathbf{E}_g$ containing the average values of $n_g$ individuals in $g$.

In order to map from OCEAN to emotion parameters we observe some aspects in literature [14]:

- O− : person is close to interact with others;
- O+ : person is aware of his/her feelings;
- C+ : person is optimistic;
- C− : person is pessimist;
- E+ : person has a strong relationship with positive emotions;
- E− : person presents relationship with negative emotions;
- A+ : person has a strong relationship with positive reactions;

- A− : person presents relationship with negative reactions;
- N− : known by the emotional stability;
- N+ : person feels negative emotions;

Such data resulted in empirical definitions included in Table 2, that shows the mapping from OCEAN traits to the chosen emotions. In fact, we were not the first one to propose this type of mapping. Saifi et al. [37] proposes similar data for a different approach where authors were interested in providing critical emotions in crowd simulators. Davis and Panksepp [15] also proposed a similar approach unifying basic emotions with personality.

In Table 2, the plus/minus signals along each factor represent the positive/negative value of each one. For example, O+ stands for positive values (i.e., $O \geq 0.5$) and O− stands for negative values (i.e., $O < 0.5$)). A positive value for a given factor (i.e., 1) means the stronger the OCEAN trait is, the stronger is the emotion too. A negative value (i.e., $-1$) does the opposite, therefore, the stronger the factor's value, the weaker is a given emotion. A zero value means that a given emotion is not affected at all by the given factor. To better illustrate, a hypothetical example is given: If an indi-

**Table 2** Emotion mapping from OCEAN to OCC

| Factor | Fear | Happiness | Sadness | Anger |
|---|---|---|---|---|
| O+ | 0 | 0 | 0 | −1 |
| O− | 0 | 0 | 0 | 1 |
| C+ | −1 | 0 | 0 | 0 |
| C− | 1 | 0 | 0 | 0 |
| E+ | −1 | 1 | −1 | −1 |
| E− | 1 | 0 | 0 | 0 |
| A+ | 0 | 0 | 0 | −1 |
| A− | 0 | 0 | 0 | 1 |
| N+ | 1 | −1 | 1 | 1 |
| N− | −1 | 1 | −1 | −1 |

The plus/minus signals along each factor represent the positive/negative value of each one



**Fig. 4** OCEAN representation of each individual in the scene. Each color bar is related to one OCEAN dimension: blue is *Openness*, cyan is *Conscientiousness*, green is *Extraversion*, yellow is Agreeableness and red, *Neuroticism*. Information about highlighted group *G*1 (formed by the individuals *P*9 and *P*10) is detailed in Fig. 5 (color figure online)

vidual has a high value for extraversion (for example, E = 0.9), following the mapping in Table 2, this individual can present signals of happiness (i.e., If E+ then Happiness = 1) and should not be angry (i.e., If E+ then Anger= −1).

Finally, emotions are normalized in the interval [0; 1] considering the lowest and highest achieved values in the video, in order to keep the obtained scale.
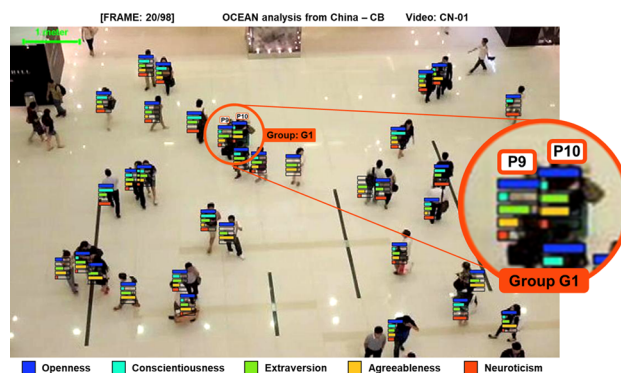
## 4 Experimental results

In this section we discuss some results obtained with our approach. We organized it into four different analysis: (i) OCEAN and Emotion recognition in spontaneous videos (Sects. 4.1 and 4.2, respectively), (ii) an OCEAN and Emotion comparison with Literature (Sect. 4.3), (iii) an OCEAN and Emotion analysis in the Fundamental Diagram experiment (Sect. 4.4) and (iv) a quantitative analysis of personal space (Sect. 4.5).

### 4.1 OCEAN analysis

In this section, we present OCEAN analysis involving spontaneous videos (crowds in public spaces). Firstly, we calculate the OCEAN of each individual in the video at each frame. Once we get the individuals OCEAN, the group OCEAN is computed by the average of the individuals' OCEAN that are part of the group. Figure 4 shows a representation of each individual OCEAN in a determined frame. We used five color box that represent the five dimensions, where blue is related to openness, cyan indicates conscientiousness, green indicates extraversion, yellow means agreeableness, and red indicates neuroticism.

The group of individuals highlighted in Fig. 4 is composed of two people. The OCEAN of this group (named *G*1) over the time is illustrated in Fig. 5a, and obtained by the aver-
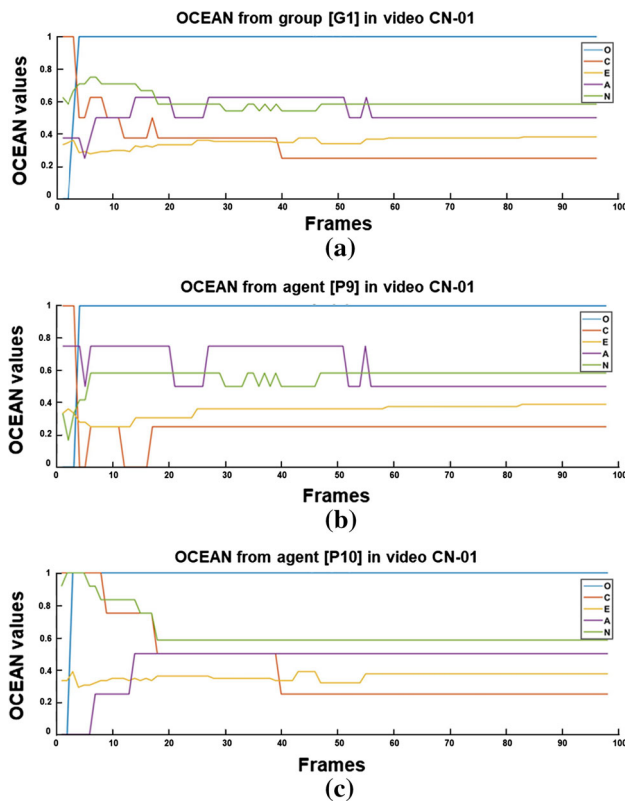
age OCEAN of its individuals (*P*9 and *P*10), presented in Fig. 5b, c. As can be seen in such figures, the two individuals present more motion variation at the beginning of the video and then they keep the same motion characteristics until the end of the short movie (which duration is 100 frames). Their Openness is high because they keep low angular variation along their trajectories. On the other hand, their Conscientiousness dimension is lower than other dimensions because they keep low speeds in comparison to other groups.

The group OCEAN reflects the individuals OCEAN. An analysis of OCEAN in the country level is presented later in this section. Once we have the OCEAN values, individuals emotions are detected in the analyzed videos, as discussed in next section.

### 4.2 Emotion analysis

This section presents the emotion analysis in spontaneous videos. Figure 6 shows an example of the emotion detection in a video from Austria. A filled square represents that the person has a positive value from that emotion, a half filled square means that the emotion is neutral and a not filled square means that the person has a negative value from the emotion. For example, the highlighted person, with the blue arrow, got a negative value regarding the Anger emotion (the red square is not filled). It happens because this individual is interacting with the other (walking in the same direction and with similar angular variation), so collectivity and socialization levels are high and isolation level is low, consequently Anger receives a negative score.

Still in Fig. 6, the highlighted person with the orange arrow gets a zero value for the Happiness emotion (the green square is half filled). It happens because she/he is alone, changing orientation (angular variation), with high isolation. The person highlighted with the red arrow gets

**Fig. 5** OCEAN data over the time observed in the group of individuals highlighted in Fig. 4: data from group $G1$ (**a**) and its individuals $P9$ (**b**) and $P10$ (**c**)
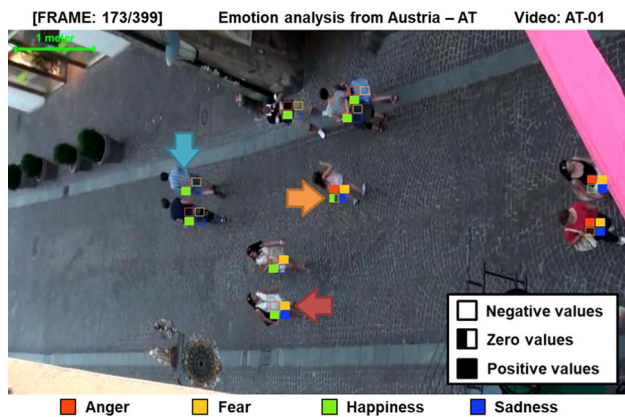


**Fig. 6** Emotion representation based on the OCEAN mapping process. Each square represents one emotion, where a filled square represents that the person has a positive value from an emotion, a half filled square means that the emotion is neutral and a not filled square means that the person has a negative value from an emotion. Each color is related to one emotion: a red square is related to *Anger*, an yellow square means *Fear*, green indicates *Happiness* and blue, *Sadness* (color figure online)

a positive value for the Fear (the yellow square is completely filled). It happens because this individual is moving slower and with high angular variation, so his/her dimension C is low, generating a high value for Fear. The legend
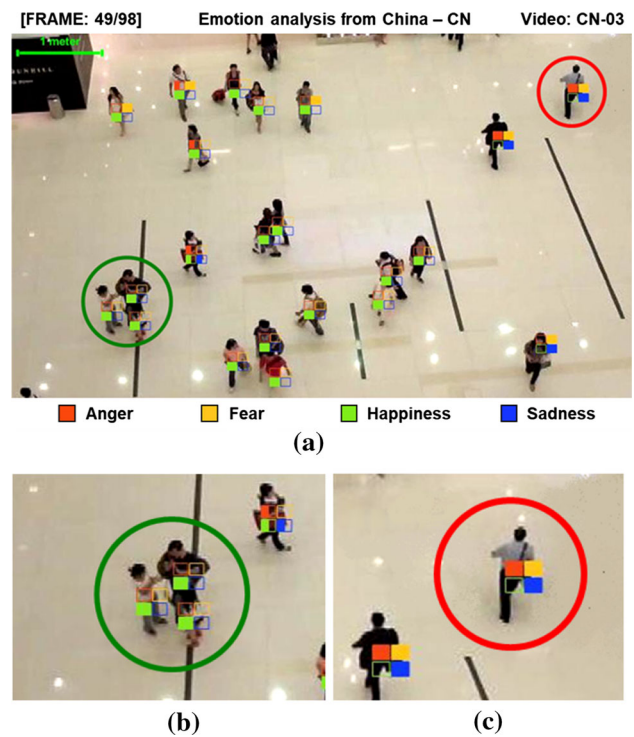


**Fig. 7** Emotion analysis in the video $CH - 03$ from China: emotion of each individual in the scene (**a**), focused on a group (**b**) and one individual (**c**) highlighted. Based on our approach, individuals in the group (**b**) tend to be happy, while the individual alone (**c**) tends to experience negative emotions

of colors is the following: red is related to anger, yellow means fear, green indicates happiness, and blue indicates sadness.

In another example, showed in Fig. 7a, we highlight two different situations, a group (green circle) and an individual alone (red circle). It is interesting to notice that individuals who are part of a bigger group or have a high collectivity tend to be happy, as we can see in the highlighted group in Fig. 7b. On the other hand, individuals who are alone and distant from others tend to experience negative emotions (see an example in Fig. 7c).

### 4.3 OCEAN and emotion comparison with literature

In the next experiment we compare our results with available literature [14] about OCEAN values of different Countries. In the case of the present work we focus on Germany and Brazil, since in next section we also present a comparative about these two Countries. It is important to emphasize that the data registered in [14] was acquired based on subject answers on surveys and not based on their visual behavior in video sequences. We use OCEAN dimensions of the two analyzed Countries (Brazil and Germany) as presented by Costa et al. [14] as ground-truth in our approach. We evaluated our
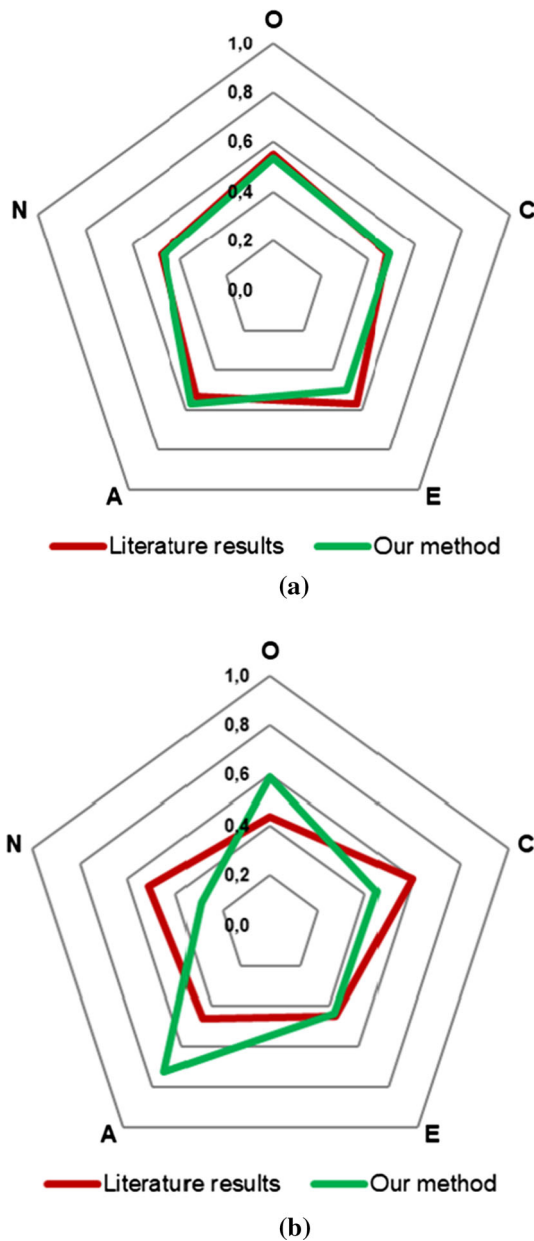
(a)



(b)

**Fig. 8** OCEAN comparison among our approach and literature [14] values: **a** data from Brazil and **b** data from Germany

method in a set of 10 available videos, where 2 videos were from Germany and 8 from Brazil.

We evaluate the accuracy of our approach to detect the OCEAN values of the Countries based on percentage difference when compared with the literature results, considering all dimensions among all videos. Figure 8 shows the results obtained with our approach for two countries in all OCEAN dimensions, in comparison with the literature [14]. It is interesting to highlight that results achieved for Brazil showed a higher accuracy, when compared to Germany. In the same time, this was the country with more videos to be analyzed.
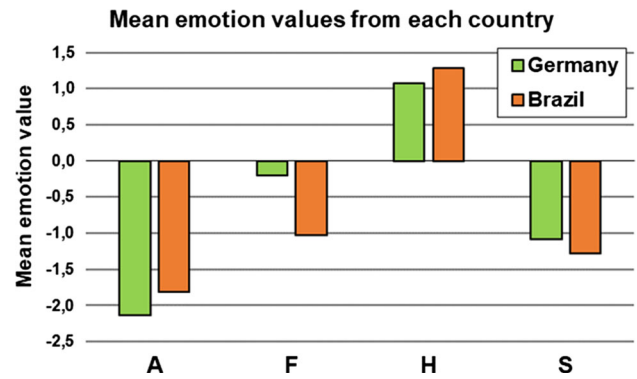


**Fig. 9** Mean emotion from each country, considering all the 10 videos from the experiments (two from Germany and 8 from Brazil)
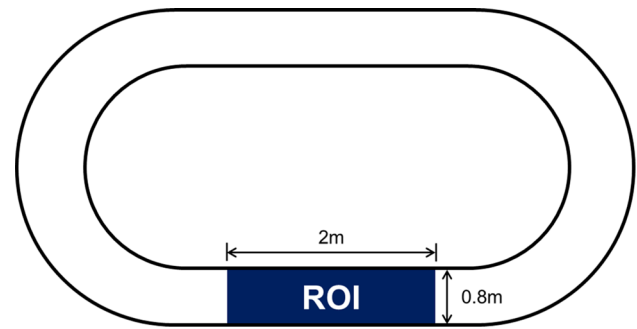


**Fig. 10** Sketch of experimental setup according to [11]. The Region Of Interest (ROI) is a rectangle of $2 \times 0.8$ m

In the same way, we compute the emotions in the country level. For this, similar to the OCEAN approach, the emotion of each country is computed as the mean emotion from the videos from that country. Figure 9 shows the mean emotion from the analyzed countries in this experiment. As far as we know, there is no data to be compared in this matter.

### 4.4 OCEAN and emotion analysis in the fundamental diagram experiment

In last sections we were interested about detecting personality and emotion traits in videos from different Countries, in public spaces. However, since the contexts are not the same, e.g., people can be close or faraway from others not just because their personality but because the contexts they are evolving on or the relationship with spaces are not the same. Consequently, this research presents such difficult challenge to solve in order to evaluate/validate results. The ideal is to discard the spatial context in order to analyze only the people behavior, from different Countries, while executing the same task. Therefore, we found the Fundamental Diagram applied to pedestrians, as proposed in [11]. We detect personality and emotion traits in video sequences from two
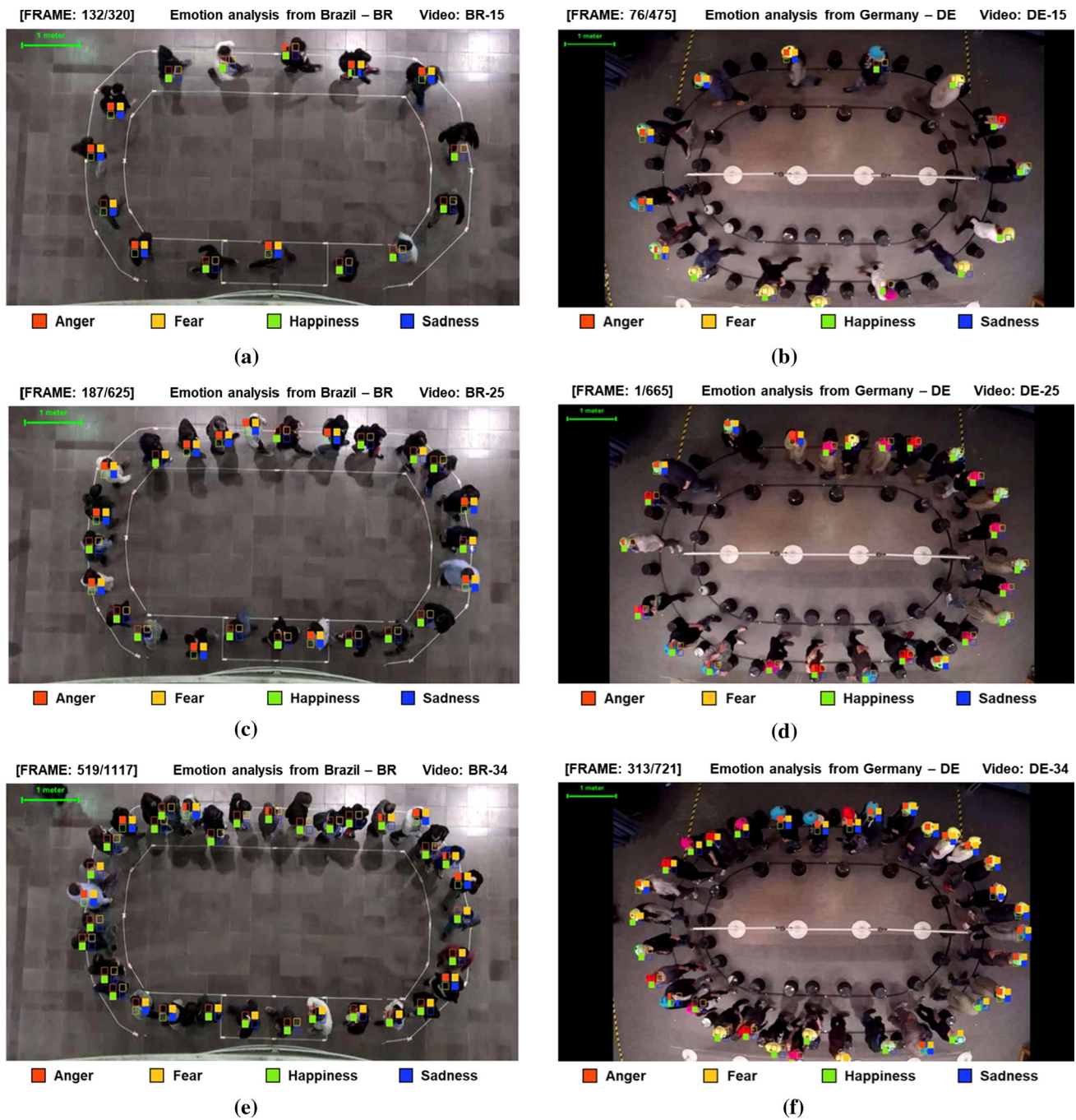
**Fig. 11** Emotion analysis from the individuals performing the experiment. Left side: Experiment from Brazil, with 15 (**a**), 25 (**c**) and 34 individuals (**e**). Right side: Experiment from Germany, with 15 (**b**), 25 (**d**) and 34 individuals (**f**)

countries (Brazil and Germany[2]), where all the individuals are performing exactly the same task, i.e., walking in a controlled space. The experiments have been performed in an environment setup as illustrated in Fig. 10.

This experiment was conducted as described in [11], being the same in all the two countries, with the same populations ($N = 15$, 25 and 34). The corridor was built up with markers and tape on the ground. Its size and shape is presented in Fig. 10. The length of the corridor is 17.3 m, while the width of the passageway is $0.8m$, which is sufficient for a single person walk. In addition, we can observe on the bottom of Fig. 10 a rectangle of 2 x 0.8 meters which illustrates the

---

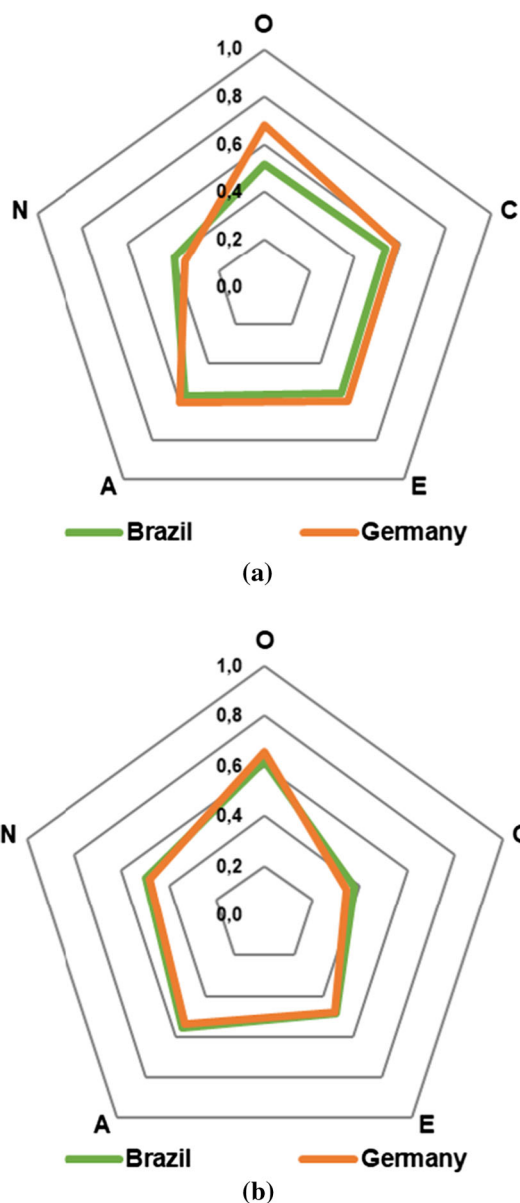[2] We have access to such videos thanks to the authors of database of PED experiments, available at http://ped.fz-juelich.de/db/.

**(a)**



**(b)**

**Fig. 12** Averaged OCEAN values observed in each country: **a** when $N = 15$ individuals and **b** when $N = 34$ individuals



**Fig. 13** Pearson's correlation of OCEAN values obtained in videos from Germany and Brazil, with different populations ($N = 15$, $N = 25$ and $N = 34$)



**Fig. 14** Pearson's correlation of emotion values obtained in videos from Germany and Brazil, with different populations ($N = 15$, $N = 25$ and $N = 34$)

Region of Interest (ROI) where the populations were captured to be analyzed, as proposed in [11].

Figure 11 shows examples of emotion detection performed by individuals in the experiment of Fundamental Diagram. Pictures show the performed tests with different sizes of population in both Countries.

Figure 12 shows the OCEAN values from Brazil and Germany in the experiments with extreme sizes of populations, i.e., $N = 15$ and $N = 34$. Based only on a visual inspection, we can easily perceive that OCEAN values from the both Countries are more similar when the density of people is higher in the experiment. It can indicate that people assumes group-level behavior instead of individual-level
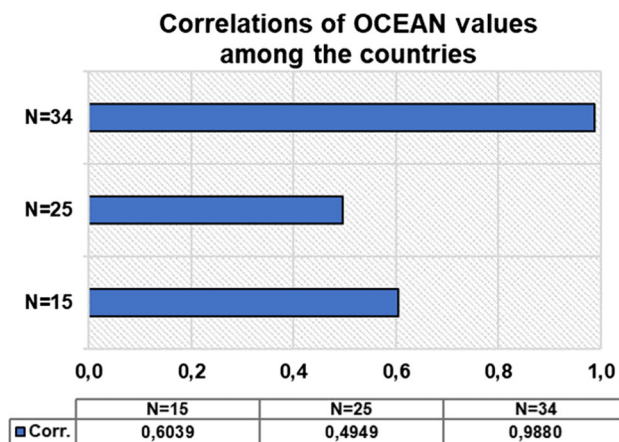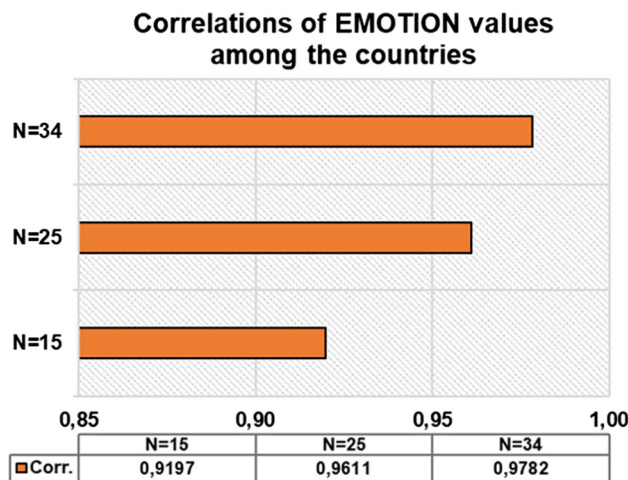
behavior caused by the higher density and the lack of free space. It agrees with several theories about mass behavior as discussed in [4,43].

In this analysis, we also compute Pearson's correlation to find out the similarity between the two countries in both OCEAN and emotion aspects. Figure 13 shows the Pearson's correlation of OCEAN values among the countries, while Fig. 14 shows the Pearson's correlation of the emotion values.

Considering the both Figs. 13 and 14, it is possible to see that, except from the population $n = 25$ in OCEAN experiment, as the density of people in the experiment increases, the correlation among the countries also increases. We also included an investigation regarding the emotion and OCEAN impacted by the density of people. This was possible because the performed task is the same (i.e., individuals are walking in the same predefined environment) while only the number of
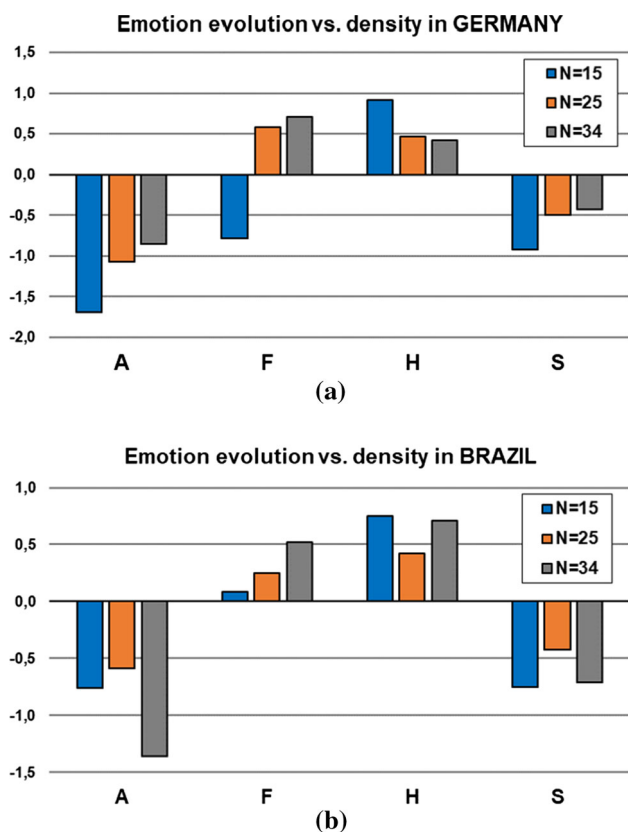
Fig. 15 Impact of emotion on the variation of densities ($N = 15$, $N = 25$ and $N = 34$) in both countries: **a** Germany and **b** Brazil



Fig. 16 Preferred distances observed in our approach versus Sorokowska [42] in different populations ($N = 15$, $N = 20$, $N = 25$ and $N = 30$)

people increases. Figure 15a shows how the emotion values vary according to the density of people in videos from Germany. It is interesting to see how the emotions Anger, Fear and Sadness increases proportionally as the density increases too. On the other hand, Happiness emotion decreases proportionally as a function of observed density. Indeed, the data observed in Germany was in accordance to what was empirically expected in our hypothesis, i.e., the only positive emotion (H) decreases as the density increases. However, the computed emotion for Brazil was not so well behaved as in Germany, and certainly it could be better investigated in a future work. One possible explanation is that Brazilian people were colleagues/friends while in Germany, they were related in [11] as volunteers to the experiment.

Although we did not find any information about emotion and personality detection in videos from different Countries related in literature, these results seem promising in order to understand people personality, emotion and cultural differences in video sequences.

### 4.5 Quantitative analysis: personal space

In this experiment we performed a comparison among the preferred distance people keep from others, as evaluated in
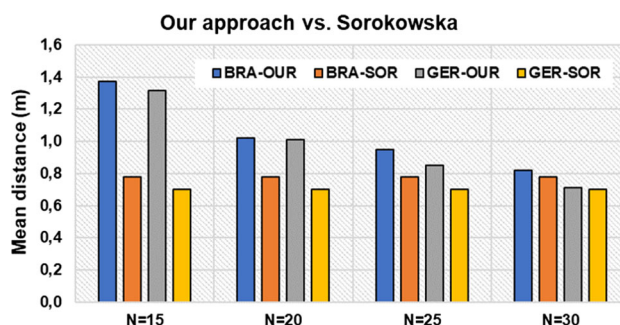
a study performed by [42]. Results obtained with the experiment performed in our approach are compared with results of [42], as shown in Fig. 16. In the Sorokowska work, the answers were given on a distance (0–220 cm) scale anchored by two human-like figures, labeled A and B. Participants were asked to imagine that he or she is Person A. The participant was asked to rate how close a Person B could approach, so that he or she would feel comfortable in a conversation with Person B.

In our approach we measure the distances a person A keeps from a person B right in front of he or she, in the video sequences. For the comparison, in our approach, we use the distances from the Fundamental Diagram video sequences varying the people number ($N = 15$, $N = 20$, $N = 25$ and $N = 30$) and from the Sorokowska's approach we select the evaluation from acquaintance people, where the people are not close neither strangers, similar to people in our experiment. As we can see in Fig. 16, in spite of the fact that distances from our approach are higher than the ones from Sorokowska, when there are few people ($N = 15$ and $N = 20$), the proportion is similar in both scenarios. In addition, in the experiment $N = 30$, the distances observed in our approach are close to those obtained in the Sorokowska study. In all scenarios, people from Brazil keeps higher distances from others than people from Germany. The same happens in Sorokowska's research. Although they are different experiments (videos and surveys), our method indicates that there is a correlation between people feeling in abstracted environment (as in surveys) and as they behave in video sequences.

## 5 Final considerations

In this paper, we describe a way to detect individual-level traits of emotion and personality observed in video sequences, based on individuals and groups features. We propose to detect OCEAN personality traits and compared with data from different countries existent in the literature. In

addition, based on OCEAN we compute 4 traits of emotion defined in OCC model. We believe the results are promising and video sequences can be used to detect personality and emotion, what can help us to understand people behavior in video sequences.

An important challenge in this area is the comparison with real life data. In this work we successfully compared our results with OCEAN from two specific countries (Brazil and Germany) present in Psychological literature. As one particular aspect to be considered in behavior analysis is the context and environment in which individuals behave, we decided to exclude such variations by fixing tasks that the tested populations were required to execute. This is why we used Fundamental Diagram for pedestrians. To do that we performed a full experiment in Brazil to serve as benchmark to our research. Therefore, we measured the preferred personal distance from individuals in FD in order to present a quantitative data analysis.

The results obtained from our approach indicate that our model generates coherent information when compared to data provided in available literature, as shown in various analysis. It is important to note that the mapping to OCEAN and EMOTION dimensions was empirically defined through equations using data extracted from computer vision. NEO PI-R results in the literature measured these dimensions by considering a different type of information (subjective responses of individuals collected through questionnaires). The results of preferred distances presented by Sorokowska et al. [42] also were based on subjective responses informed by individuals.

For our future work, we consider to select items related to collectiveness from the International Personality Item Pool (http://www.ipip.ori.org/), which originated the NEO PI-R, to establish a questionnaire that can identify individual-level traits that correspond to group behaviors and could be analyzed computationally. One way to do this would be to ask participants to answer the proposed questionnaire, then divide these participants into groups with high and low scores, and, finally, ask these groups to perform some behavioral tasks, tasks that could be mapped in videos of crowds. Such a set of individual-level traits related to group behavior will not only increase the accuracy of group analysis tools by means of videos as well as group simulations based on the unique characteristics of each agent.

## References

1. Baig, M.W., Baig, M.S., Bastani, V., Barakova, E.I., Marcenaro, L., Regazzoni, C.S., Rauterberg, M.: Perception of emotions from crowd dynamics. In: 2015 IEEE International Conference on Digital Signal Processing (DSP), pp. 703–707. IEEE (2015)
2. Barry, B., Stewart, G.L.: Composition, process, and performance in self-managed groups: the role of personality. J. Appl. Psychol. **82**, 62 (1997)
3. Bins, J., Dihl, L.L., Jung, C.R.: Target tracking using multiple patches and weighted vector median filters. MIV **45**(3), 293–307 (2013). https://doi.org/10.1007/s10851-012-0354-y
4. Bon, G.: The Crowd: A Study of the Popular Mind. Cosimo Classics Personal Development. Cosimo Classics, New York (2006)
5. Cai, Z., Yu, Z.L., Liu, H., Zhang, K.: Counting people in crowded scenes by video analyzing. In: 9th IEEE ICIEA, pp. 1841–1845 (2014). https://doi.org/10.1109/ICIEA.2014.6931467
6. Cattell, R.B.: The primary personality factors in women compared with those in men. Br. J. Stat. Psychol. **1**(2), 114–130 (1948). https://doi.org/10.1111/j.2044-8317.1948.tb00231.x
7. Cattell, R.B.: Personality: A Systematic, Theoretical, and Factual Study, 1st edn. McGraw-Hill, New York (1950)
8. Chan, A.B., Vasconcelos, N.: Bayesian poisson regression for crowd counting. In: 12th IEEE ICCV, pp. 545–551 (2009)
9. Chandran, A., Poh, L.A., Vadakkepat, P.: Identifying social groups in pedestrian crowd videos. In: ICAPR, pp. 1–6 (2015). https://doi.org/10.1109/ICAPR.2015.7050677
10. Chattaraj, U., Seyfried, A., Chakroborty, P.: Comparison of pedestrian fundamental diagram across cultures. ACS **12**(03), 393–405 (2009). https://doi.org/10.1142/S0219525909002209
11. Chattaraj, U., Seyfried, A., Chakroborty, P.: Comparison of pedestrian fundamental diagram across cultures. ACS (2009). https://doi.org/10.1142/S0219525909002209
12. Clore, G.L., Ortony, A.: Psychological construction in the OCC model of emotion. Emot. Rev. **5**(4), 335–343 (2013)
13. Costa, Jr. P.T., McCrae, R.R.: Revised NEO Personality Inventory (NEO-PI-R) and the NEO Five-Factor Inventory (NEO-FFI) Professional Manual. Psychological Assessment Resources, Odessa (1992)
14. Costa, Jr. P.T., McCrae, R.R.: NEO PI-R: inventrio de personalidade NEO revisado. Sao Paulo (2007)
15. Davis, K.L., Panksepp, J.: The brain's emotional foundations of human personality and the affective neuroscience personality scales. Neurosci. Biobehav. Rev. **35**(9), 1946–1958 (2011)
16. Davis, K.L., Panksepp, J., Normansell, L.: The affective neuroscience personality scales: normative data and implications. Neuropsychoanalysis **5**(1), 57–69 (2003)
17. Digman, J.M.: Personality structure: emergence of the five-factor model. Annu. Rev. Psychol. **41**(1), 417–440 (1990). https://doi.org/10.1146/annurev.ps.41.020190.002221
18. Durupinar, F., Allbeck, J., Pelechano, N., Badler, N.: Creating crowd variation with the ocean personality model. In: Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems—vol. 3, pp. 1217–1220. IFAAMAS, Richland (2008). http://dl.acm.org/citation.cfm?id=1402821.1402835
19. Ekman, P., Friesen, W.V.: Constants across cultures in the face and emotion. J. Pers. Soc. Psychol. **17**(2), 124 (1971)
20. Favaretto, R.M., Dihl, L., Musse, S.R.: Detecting crowd features in video sequences. In: Proceedings of Conference on Graphics, Patterns and Images (SIBGRAPI). IEEE Computer Society Conference Publishing Services (2016). https://doi.org/10.1109/SIBGRAPI.2016.34
21. Favaretto, R.M., Dihl, L., Musse, S.R., Vilanova, F., Costa, A.B.: Using big five personality model to detect cultural aspects in crowds. In: 2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), pp. 223–229 (2017). https://doi.org/10.1109/SIBGRAPI.2017.36
22. Feng, L., Bhanu, B.: Understanding dynamic social grouping behaviors of pedestrians. IEEE STSP **9**(2), 317–329 (2015)
23. Fiske, D.W.: Consistency of The Factorial Structures in Personality Ratings from Different Sources. University of MICHIGAN (1948). https://books.google.com.br/books?id=o74lnQEACAAJ
24. Gorbova, J., Lsi, I., Litvin, A., Anbarjafari, G.: Automated screening of job candidate based on multimodal video processing. In:

2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pp. 1679–1685 (2017). https://doi.org/10.1109/CVPRW.2017.214

25. Guy, S.J., Kim, S., Lin, M.C., Manocha, D.: Simulating heterogeneous crowd behaviors using personality trait theory. In: Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, SCA '11, pp. 43–52. ACM, New York (2011). https://doi.org/10.1145/2019406.2019413

26. Hall, E.: The Hidden Dimension. A Doubleday Anchor Book. Anchor Books (1990). http://books.google.com.br/books?id=zGYPwLj2dCoC

27. Hall, C.S., Lindzey, G., Campbell, J.B.: Theories of Personality, 4th edn. Wiley, New Jersey (1998)

28. Jo, H., Chug, K., Sethi, R.J.: A review of physics-based methods for group and crowd analysis in computer vision. J. Postdr. Res. 1(1), 4–7 (2013)

29. Kaup, D., Clarke, T., Malone, L., Oleson, R.: Society for computer simulation. Simul. Ser. 38(4), 365–370 (2006)

30. Lord, W.: Neo Pi-R A Guide to Interpretation and Feedback in a Work Context, 1st edn. Hogrefe Ltd, Oxford (2007)

31. McCrae, R., Terracciano, A.: Universal features of personality traits from the observer's perspective : data from 50 cultures. J. Pers. Soc. Psychol. 88(3), 547–561 (2005)

32. McCrae, R.R.: NEO-PI-R Data from 36 Cultures, pp. 105–125. Springer US, Boston (2002). https://doi.org/10.1007/978-1-4615-0763-5_6

33. Montag, C., Panksepp, J.: Primary emotional systems and personality: an evolutionary perspective. Front. Psychol. 8, 464 (2017)

34. Ortony, A., Clore, G.L., Collins, A.: The Cognitive Structure of Emotions. Cambridge University Press, Cambridge (1990)

35. Rabiee, H., Haddadnia, J., Mousavi, H., Nabi, M., Murino, V., Sebe, N.: Emotion-based crowd representation for abnormality detection. arXiv preprint arXiv:1607.07646 (2016)

36. Revelle, W., Scherer, K.R.: Personality and emotion. In: Sander D, Scherer K (eds) Oxford Companion to Emotion and the Affective Sciences. Oxford University Press, New York, pp. 304–306 (2009)

37. Saifi, L., Boubetra, A., Nouioua, F.: An approach for emotions and behavior modeling in a crowd in the presence of rare events. Adapt. Behav. 24(6), 428–445 (2016)

38. Sethi, R.J.: Towards defining groups and crowds in video using the atomic group actions dataset. In: 2015 IEEE International Conference on Image Processing, ICIP 2015, Quebec City, QC, Canada, September 27-30, 2015, pp. 2925–2929 (2015). https://doi.org/10.1109/ICIP.2015.7351338

39. Shao, J., Loy, C.C., Wang, X.: Scene-independent group profiling in crowd. In: IEEE CVPR, pp. 2227–2234 (2014). https://doi.org/10.1109/CVPR.2014.285

40. Solera, F., Calderara, S., Cucchiara, R.: Structured learning for detection of social groups in crowd. In: 10th IEEE AVSS (2013)

41. Solmaz, B., Moore, B.E., Shah, M.: Identifying behaviors in crowd scenes using stability analysis for dynamical systems. IEEE PAMI 34(10), 2064–2070 (2012). https://doi.org/10.1109/TPAMI.2012.123

42. Sorokowska, A., Sorokowski, P.: P.H.: preferred interpersonal distances: a global comparison. J. Cross Cult. Psychol. 48(4), 577–592 (2017). https://doi.org/10.1177/0022022117698039

43. Vilanova, F., Beria, F., et al.: A.C.: deindividuation: from le bon to the social identity model of deindividuation effects. CP 4(1) (2017). https://doi.org/10.1080/23311908.2017.1308104

44. Zhang, Y., Qin, L., Ji, R., Zhao, S., Huang, Q., Luo, J.: Exploring coherent motion patterns via structured trajectory learning for crowd mood modeling. IEEE Trans. Circuits Syst. Video Technol. 27(3), 635–648 (2017)

45. Zhou, B., Tang, X., Zhang, H., Wang, X.: Measuring crowd collectiveness. IEEE PAMI 36(8), 1586–1599 (2014). https://doi.org/10.1109/TPAMI.2014.2300484

**Rodolfo Migon Favaretto** is currently a Ph.D. candidate in the Department of Computer Science at Pontifical Catholic University of Rio Grande do Sul (PUCRS) and also a professor of the Information and Communication Department in Federal Institute Sul-rio-grandense (IFSul) since 2015. He received his master's degree in Computer Science from the Federal University of Pelotas (UFPel). He graduated in Computer Science from Universidade Regional Integrada e das Missões (URI).

**Paulo Knob** holds a degree in Computer Science from the University of Passo Fundo (2010), with specialization in Information Security by SENAC (2014), holds a master's in Computer Science from by PUCRS (2018), and currently is a Ph.D. student at PUCRS, led by Professor Soraia Raupp Musse.

**Soraia Raupp Musse** received Ph.D. degree in Computer Science from the École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland, in 2000, under the supervision of Prof. D. Thalmann. She is currently an Associate Professor with the Department of Computer Science, Pontifical Catholic University of Rio Grande do Sul, Porto Alegre, Brazil, where she created (and still coordinates) the Virtual Human Laboratory, where projects supported by private companies and the Brazilian government are developed. Her current research interests include crowd simulation and analysis, facial animation and integration between computer graphics, pattern recognition, and computer vision.

**Felipe Vilanova** is an Academician at the Universidade Federal do Rio Grande do Sul (UFRGS) and a Member of the Center for psychological studies on at-risk populations (Cep-Rua, UFRGS) and the research laboratory Prejudice, Vulnerability and Psychosocial Processes (PVPP, PUCRS).

**Ângelo Brandelli Costa** is a Professor of the Psychology and Social Sciences Graduate Programs of the Pontifical Catholic University of Rio Grande do Sul (PUCRS), is graduated in Psychology from the Federal University of Rio Grande do Sul (UFRGS), is a specialist in Social Psychology and Health Psychology, and has a master's in Social and Institutional Psychology (UFRGS) and PhD in Psychology (UFRGS).