

Emotion, Personality and Cultural Aspects in Crowds: towards a Geometrical Mind

Rodolfo Migon Favaretto¹

Virtual Humans Simulations Laboratory - VHLab
Pontifical Catholic University of Rio Grande do Sul
Email: rodolfo.favaretto@acad.pucrs.br

Soraia Raupp Musse

Virtual Humans Simulations Laboratory - VHLab
Pontifical Catholic University of Rio Grande do Sul
Email: soraia.musse@pucrs.br

Abstract—In this work we proposed a computational model to extract pedestrian characteristics from video sequences. The proposed model considers a series of characteristics of the pedestrians and the crowd, such as number and size of groups, distances, speeds, among others, and performs the mapping of these characteristics in personalities, emotions and cultural aspects, considering the Cultural Dimensions of Hofstede (HCD), the Big-Five Personality Model (OCEAN) and the OCC Emotional Model. The main hypothesis is that there is a relationship between so-called intrinsic human variables (such as emotion) and the way people behave in space and time. As one of the main contributions, four large dimensions of geometric characteristics (Big4GD) were proposed: *I - Physical*, *II - Personal and Emotional*, *III - Social* and *IV - Cultural*, which seek to describe the behavior of pedestrians and groups in the crowd. The *GeoMind* tool was developed for the purpose of detecting the four geometric dimensions from video sequences. In addition, several analyzes were carried out with the purpose of validating the proposed model, from comparing results with the literature, including the comparison of spontaneous multitudes from several countries and controlled experiments involving Fundamental Diagrams.

I. INTRODUCTION

The study of human behavior is a subject of great scientific interest and probably an inexhaustible source of research [1]. There are many applications such as entertainment (games and movies), understanding of human behavior, security and surveillance, urban planning and activity recognition [2], [3]. When a group of individuals shares the same physical space and has a common goal, they have the characterization of a collective and highly dynamic or, more broadly, a crowd [4].

Chattaraj et al. [5] suggest that cultural and population differences can produce changes in the speed, density, and flow of the crowd. Interested in these aspects, Geert Hofstede created a model of culturality, which became known as the Cultural Dimensions of Hofstede (HCD) [6]. Another factor that can interfere in the cultural aspects of a group of individuals are the personality and emotion of each one. Regarding personality, we adopted the personality model *Big-Five* [7], [8]. It is a descriptive (taxonomic) psychological model of five factors of personality traits. Related to emotions, Ortony, Clore and Collins [9] proposed an emotion model called OCC.

Following this line of research, this work seeks to explore the cultural factors of multi-country crowds from video sequences. The main goal is to **develop a computational**

model for detection and analysis of pedestrian and group features in crowds from video sequences. The idea is to use characteristics of the crowd and their individuals, such as: average distance between people, number and size of groups, speed, density, directions, among other factors that may vary according to the type and nationality of the population to detect and characterize personal, social, emotional and cultural aspects. In particular, we want to propose new factors/dimensions that can be used to characterize crowds, groups and their individuals relating intrinsic features as cultural aspects, personalities and emotions to physical and geometrical manifestations (distances, speeds and so on).

The relevance of this research is justified due its diverse applications, such as physical space planning, entertainment and security management. These areas seek to consider regional and cultural aspects, but usually this is done with empirical knowledge, there is no computational methods to help in this task. In order to address this problem, we proposed the development of a computational model that allows the extraction of pedestrian and crowd characteristics based on their behaviours manifestation in space/time, being able to identify cultural, personality and emotional aspects and behaviors that differ from one country to another. Figure 1 shows an overview of the features mapping from the proposed approach.

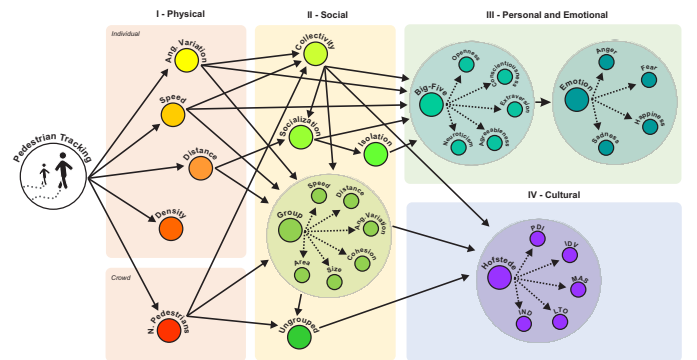


Fig. 1. **Our Big-Four Geometrical Dimensions:** features mapping from the proposed approach divided into four dimensions: *I-Physical*, *II-Social*, *III-Personal and Emotional* and finally, *IV-Cultural*.

As one of the main contributions of this thesis, we proposed four different dimensions of features to characterize pedestrians organized in groups/crowds in video sequences. These

¹This work relates to a Ph.D thesis.

dimensions, named as **Big-Four Geometrical Dimensions** or just **Big4GD** model, illustrated in Figure 1, are:

- *I – Physical*, which keeps the physical features of pedestrians obtained directly from the tracking, such as speeds, distances from a pedestrian from others, angular variations, among others;
- *II – Social*, which derives from the Physical dimension and deals with social interaction, characterizing groups of pedestrians and social features, as collectivity, isolation and socialization levels of individuals;
- *III – Personal and Emotional*, which maintains the features related to personality (Big-Five) and emotion (OCC) traits;
- *IV – Cultural*, which deals with features regarding cultural aspects, according to Hofstede (HCD).

Our main hypothesis is that it is plausible to propose a computational model that based on pure geometrical input information can be used to find out some intrinsic information, i.e. correlate geometric manifestations with emotion/personality/cultural aspects. A comment about the contribution of this work: we did not find any model in literature that focus on find out and characterizing emotion, personalities and cultural aspects of people in pedestrians from video sequences related to their spatial/time behaviours.

II. A STATE-OF-THE-ART REVIEW

In the work proposed by [10], a non-recursive algorithm of motion similarity clustering (NMSC) is proposed to identify moving pedestrians together in social groups. Solera and his collaborators [11] propose a new algorithm for the detection of groups, through the clustering of trajectories, solving by means of a parametric group of correlation trained by a support vector machine (SVM). The understanding of the individual trajectories can arise from studies on the pedestrian dynamics, allowing to analyze the relation between the density of the crowd and the movement of the people. In this sense, the Fundamental Diagrams can serve as an analysis tool [12], [13]. FDs can also be used in crowd simulations [14], [15].

[16] used the FDs to analyze disasters in crowds. In the approach proposed by [15], the relationship between flow, speed and density of a real crowd is compared with simulations using the DenseSense algorithm. [17] proposed an approach to provide critical emotions in crowd simulators. The work of [18] is quite similar to the one proposed in this paper, however it uses simulation and not computer vision. The authors introduced a virtual environment that allows the creation of different types of cultural crowds which the users can interact. In the work of detecting cultural differences proposed by [5], a study was carried out to verify the cultural influences in trajectories of individuals, using the FD computed with populations from Germany and India. The work proposed by [19] presents data that aim to differentiate populations regarding their behavior of movement in crowds. The study is conducted with populations from five countries: France, Iraq, Israel, Canada and England. The work proposed by [20] aimed to investigate differences in pedestrian flows in two cultures:

Germany and China. In the next section we start to describe our approach.

III. FIRST AND SECOND DIMENSIONS: DATA EXTRACTION, CROWD TYPES AND VIDEO SIMILARITY

In this section we describe the initial steps of the *Big-Four Geometrical Dimensions* model: tracking process, data extraction and video similarity analysis. The initial detection of people in the video is performed through the real-time object detector proposed by [21]. The tracking input parameters are the initial positions of people’s heads. After the tracking step, a set of trajectories, in image coordinates, is obtained. This step can be performed by any tracker, since the tracking process is not the main contribution of this work (we used the method proposed by [22]). The next step is to correct the perspective of the image and obtain the parameters in world coordinates, for this a homographic planar projection is performed.

Once the trajectories in world coordinates are obtained, the following information is computed for each pedestrian i , at each timestep: *i*) 2D position x_i (meters); *ii*) speed s_i (meters/frame); *iii*) angular variation α_i (degrees) w.r.t. a reference vector $\vec{r} = (1, 0)$; *iv*) isolation level ϑ_i ; *v*) socialization level ϑ_i ; and *vi*) collectivity ϕ_i . To compute the collectivity affected in individual i from all n individuals, we computed:

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

and the collectivity between two individuals 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 the speed and orientation differences between two people i and j , and w_1 and w_2 are constants that should regulate the offset in meters and radians. We have used $w_1 = 1$ and $w_2 = 1$. So, values for $\varpi(i, j)$ are included in interval $0 \leq \varpi(i, j) \leq 4.34$. $\gamma = 1$ is the maximum collectivity value when $\varpi(i, j) = 0$, and $\beta = 0.3$ is empirically defined as decay constant. Hence, ϕ_i is a value in the interval $[0; 1]$.

To compute the socialization level ϑ we use an artificial neural network (ANN) with a Scaled Conjugate Gradient (SCG) algorithm in the training process to calculate the socialization ϑ_i level for each individual i . The ANN has 3 inputs (collectivity ϕ_i of person i , mean Euclidean distance from a person i to others $\bar{d}_{i,j}$ and the number of people in the Social Space¹ according to Hall’s proxemics [23] around the person n_i). In addition, the network has 10 hidden layers and 2 outputs (the probabilities of socialization and non socialization). The final accuracy was 96%. We used 16.000 samples (70% of training and 30% of validating). Once we get the socialization level ϑ_i , we compute the isolation level $\varphi_i = 1 - \vartheta_i$, that corresponds to its inverse. For more details about how this features are obtained, please refer to [24], [25]. For each individual i in a video, we computed the average for all frames and generate a vector $\vec{V}_i$ of extracted data where

¹Social space is related to 3.6 meters [23].

$\vec{V}_i = [x_i, s_i, \alpha_i, \varphi_i, \theta_i, \phi_i]$. In the next section we describe how these features are mapped into personality dimensions.

Regarding groups, to define that two pedestrians i and j belong to the same group, they have to attend to three conditions:

- **Condition 1:** if $d(\vec{X}_i, \vec{X}_j) \leq 1.2$ meters. This distance was defined based on the interpersonal space of the proxemics distances proposed by [23];
- **Condition 2:** if $\alpha(\alpha_i, \alpha_j) \leq 15^\circ$. It was empirically defined that the difference in orientation between the two pedestrians could not be greater than 15° ; and
- **Condition 3:** if $s(s_i, s_j) \leq \beta \max\{s_i, s_j\}$. The speed difference between the two pedestrians can not be greater than β , where $\beta = 5\%$ empirically defined.

At first, as soon as a group is detected, it is identified as a group *temporary*, that is, a group that is still unstable. This group becomes a *permanent* group if it maintains its formation (without the input or output of people) for at least 10% of the amount of frames of the video. After this step, a series of information about the permanent groups are computed, such as the number G^f of groups, mean speed $\bar{s}_g$ (in meters/second) among all the members of g , mean angular variation $\bar{\alpha}_g$ (in degrees) among all the members of g , group area A_g and group cohesion C_g . All these information are described in [25].

We also propose a way to find similar videos. With the goal of making a proper comparison between different videos (for example, it would not be appropriate to compare a video with some carefree people strolling in the mall with a video where there are many people at a train station, hasty, and preoccupied with the time, since the context is different). The details about that approach is described in [25]. For example, two similar videos according to our approach are the videos $v = 16$ and $v = 20$ (illustrated in Figure 2). In addition to this qualitative analysis, the similarity between these two videos is repeated in numerical data, as described in Table I.

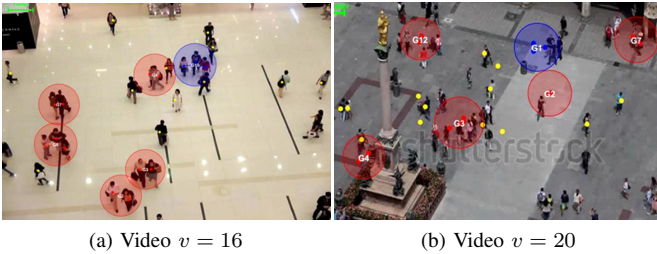


Fig. 2. **Videos 16 and 20:** red circles represent permanent groups and blue represent temporary groups. Yellow dots indicate ungrouped pedestrians.

IV. THIRD DIMENSION: DETECTION OF PERSONALITY AND EMOTION TRAITS

This section presents the proposed methodology to detect personality and basic emotion characteristics of crowds in video sequences, features from the *III - Personal and Emotional* dimension. Firstly, individuals are detected and tracked, then such information is mapped to OCEAN dimensions,

TABLE I
VIDEOS 16 AND 20: INFORMATION ABOUT SIMILARITY.

	Video $v = 16$	Video $v = 20$
Total number of pedestrians	26	28
Number of Groups	6	6
Mean pedestrians by group	2.33	2.55
Number of non-grouped pedestrians	12	13

used to find out personality and emotion in videos, based on OCC emotion model. Our model presents three main steps as following: *i)* video data extraction, *ii)* personality and *iii)* emotion analysis.

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 mapping of pedestrians, as described in Section III (using the vector $\vec{V}_i$). 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.

The five dimensions of OCEAN are: Openness (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); and Neuroticism (how much prone to psychological distress the individual is). To detect the OCEAN of each pedestrian, we used the NEO PI-R [7] that is the standard questionnaire measure of the Five Factor Model. We firstly selected NEO PI-R items related to individual-level crowd characteristics and the corresponding OCEAN-factor. For example: Like being part of crowd at sporting events corresponding to the factor Extraversion.

As we describe in details in [24], we proposed a series of empirically defined equations to map pedestrian features to OCEAN dimensions. Firstly, we selected 25 from the 240 items from NEO PI-R inventory that had a direct relationship with crowd behavior. In order to answer the items with data coming from real video sequences, we propose equations that could represent each one of the 25 items with features extracted from videos. 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 α to have answer compatible with 5. So the equation for this item is $Q_1 = s_i + \frac{1}{\alpha_i}$. In this way, we empirically defined equations for all 25 items, as presented in [24].

Figure 3 shows some results obtained by the proposed approach in all OCEAN dimensions, in comparison with the literature values presented by [26], considered as ground-truth. It is interesting to highlight that results achieved for Brazil (Figure 3(a)) showed the higher accuracy, when compared to the other countries. This was the country with more available

videos to be processed in the proposed method (9 videos), in comparison with other countries.

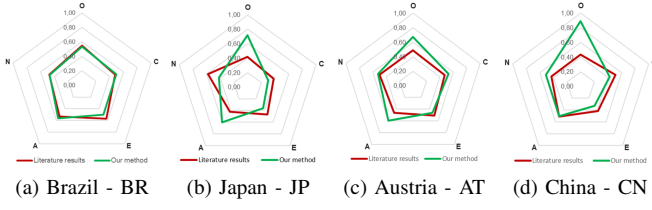


Fig. 3. **OCEAN analysis:** comparison between the proposed approach and literature values [26], from Brazil (a), Japan (b), Austria (c) and China (d).

In addition, we computed the perceptual error when accumulating each dimension from all videos and compared with literature for those Countries. With an error of 18%, the dimension *E* has the lowest value; that is an interesting observation since this was the dimension that had more questions to be analyzed. The other dimensions (*O*, *C*, *A* and *N*) presented, respectively, 46%, 26%, 37% and 22%.

Related to emotion analysis, as we presented in [27], we proposed a way to map the OCEAN dimensions of each pedestrian in OCC Emotion model. This mapping is described in Table II. In Table II, the plus/minus signals along each factor represent the positive/negative value of each one. For example concerning Openness, *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 individual has a high value for Extraversion (for example, $E = 0.9$), following the mapping in Table II, 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).

TABLE II
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

As a qualitative example, in Figure 4(a) 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 Figure 4(b). On the other hand, individuals who are alone and

distant from others tend to experience negative emotions (see an example in Figure 4(c)). It is important to keep in mind that in our approach we are not considering the internal or intrinsic emotions of pedestrians and groups, we are considering these emotions related to the physical space at a specific time.

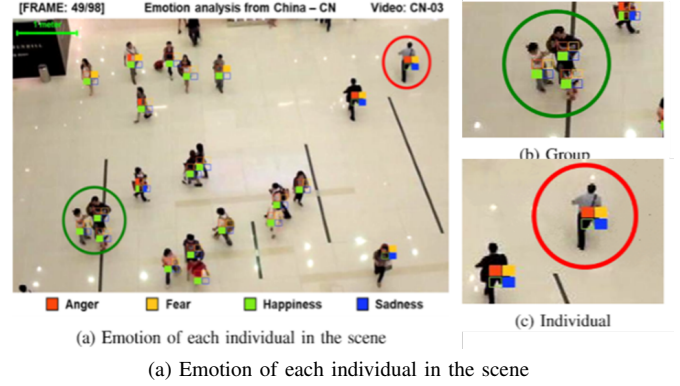


Fig. 4. **Emotion analysis in the video CN03 from China:** emotion of each individual in the scene (a), focused on a group (b) and one individual (c) highlighted. Based on our approach, individuals in the group (b) tend to be happy, while the individual alone (c) tends to experience negative emotions.

In addition, we performed an experiment to evaluate if people can perceive different personalities and emotions performed by pedestrians in crowds (all analysis are described in [28]). We used a simulation framework based on real video tracking files and replaced the pedestrians by identical avatars, generating short video sequences that were evaluated by the participants in the survey. It was interesting to see that, even without explaining to the participants the concepts of each personality or emotion and how they were calculated in our approach (considering the geometric characteristics), in all the cases, most of the participants perceived the personality and emotion that the agent was expressing in the video, in accordance with our approach.

V. FOURTH DIMENSION: DETECTION OF CULTURAL ASPECTS IN CROWDS

In this section is presented the approach to detect cultural aspects from videos considering Hofstede Cultural Dimensions [29]. In our model, the Hofstede cultural dimensions are related to the *IV - Cultural* geometrical dimension. In order to map pedestrians features in Cultural dimensions, we proposed an approach based on groups characteristics [30]. Indeed, collectivism (COL) is a % of people grouped, while the individualism (IDV) is a % of lonely people. Regarding PDI, our hypothesis is that individuals that keep close to each other recognize less the group hierarchy, while higher distances between agents can represent a more explicitly hierarchy recognition. Hence, we used the mean group distance to describe these cultural dimension (d_g). In terms of LTO/STO, the underlying idea is persistence (long-term) as opposed to quick results (short-term). So, we adapted the group orientation to this dimension, meaning that groups with higher values of angular variation result in short-term

orientation ($STO = 100 - LTO$), which are computed as shown in Equation 2.

$$LTO = \begin{cases} O_k, & \text{if } O_k \geq 50 \\ 100 - O_k, & \text{otherwise} \end{cases} \quad (2)$$

Considering the MAS dimension, we regard that the group cohesion can represent “a preference for cooperation”. So, higher levels of cohesion represent more femininity values in such dimension. Indeed, we used also LTO to weight the MAS aspect: $MAS = \sigma_1 GC_k + (1 - \sigma_1) LTO$, where $\sigma_1 = 0.5$ is the empirically chosen weight. Finally, the Indulgence vs. restraint dimension has been characterized by the groups speed and collectivism, given by $IND = \rho_1 S_k + (1 - \rho_1) COL$, where $\rho_1 = 0.5$ is an empirically chosen weight. Next section describes the developed application.

Figure 5 illustrates some results of the cultural dimensions obtained with the proposed method in comparison to the Hofstede values². It is important to note that we can not use HCD as ground truth for the proposed method, since they deal with different information (subject answers in HCD and images in the proposed method); however, it is possible to evaluate if results are coherent to be used to detect cultural aspects in different countries video sequences.

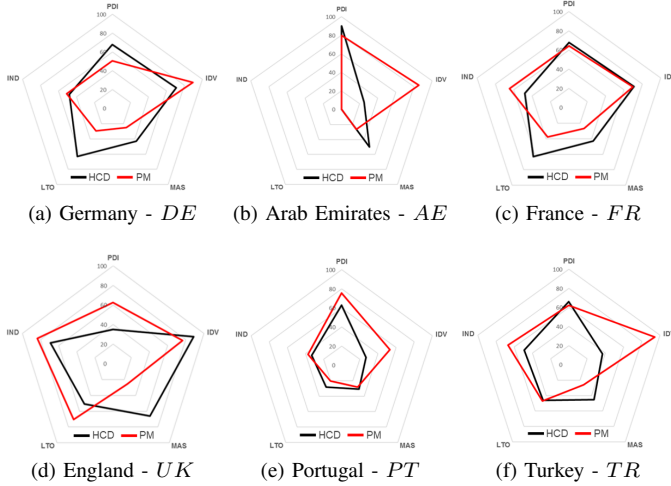


Fig. 5. **Proposed method vs. Hofstede:** results of culturality for each country. Hofstede (HCD, black line) versus the proposed method (PM, red line).

In Figure 5(b), the values of the dimensions IND and LTO on the Hofstede website (<https://geert-hofstede.com>) were not available for the country Arab Emirates, which is why the results of these two dimensions were omitted in the proposed method ($IND = 0$ and $LTO = 0$). The goal of these plots is not to use the Hofstede values as ground truth, but to verify if the results are coherent with each other, since they come from different sources (self-report in Hofstede and computational vision in the proposed method). France (FR) and Portugal (PT) presented the smallest differences in relation to the values obtained by Hofstede. The values of these countries are shown, respectively, in Figure 5(c) and (d).

²The HCD values can be found at: <http://geert-hofstede.com/>

VI. FUNDAMENTAL DIAGRAM ANALYSIS

In this section, we evaluated our model considering the Fundamental Diagrams experiment regarding personal distance and walking speeds in three different countries: India, Brazil and Germany in a controlled experiment. This experiment was applied as described in [5] to three countries with the same populations ($N=15, 20, 25, 30$ and 34). A corridor was built with markers and tape placed on the ground. Its size and shape is presented in Figure 6. At the base of the corridor was a rectangle designating the Region of Interest (ROI) from which the populations were captured as proposed by [5].

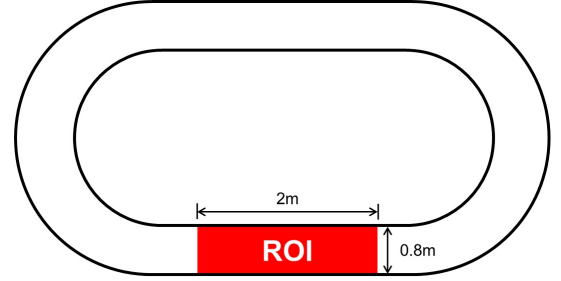


Fig. 6. **Experimental setup:** sketch of the experimental setup used as proposed by [5].

We computed the mean density and speed of individuals from the ROI per second. In turn, we determined the density (individuals per sqm) and speed (meters per second) as described by [5] and as illustrated in Figure 7 (Brazil is on the right). In this figure we plot data captured from video sequences containing 15, 20, 25, 30 and 34 individuals while walking in the experimental scenario. As we show, the Brazilian subjects travel at higher speeds than the other two samples and even at high densities (e.g., of $N = 30$ and 34).

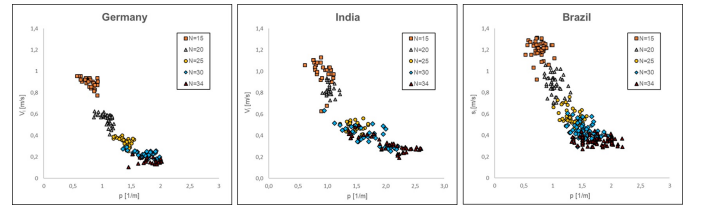


Fig. 7. **Speed density data extracted from Fundamental Diagrams for all the three countries:** (left) Germany [5], (centre) India [5] and (right) Brazil.

As for Brazilian population we had access to all of the experimental video coverage, we were able to track and determine the participants positions in each frame. However, this was not true for the populations in India and Germany, for which we have information on the time at which each person entered and exited the analysed rectangle as sent by the authors [5]. We thus used this information to compute the speed of each person in the rectangle and the distance between two consecutive individuals in the rectangle. We defined the distance from individual i to the individual in front of him/her (individual $i + 1$) as the personal space of i .

Correlations of distances between the three populations are shown in Figure 8. As can be easily observed, Pearson's correlations between the populations increase with density. Based on this, our hypothesis that high densities greatly impact personal behavioural expression makes sense as at high densities individuals act more as a mass and less as individuals [31], which ultimately affects their behaviours according to their own cultural backgrounds. This assumption is also coherent with one of the main works published on mass behaviour [32].

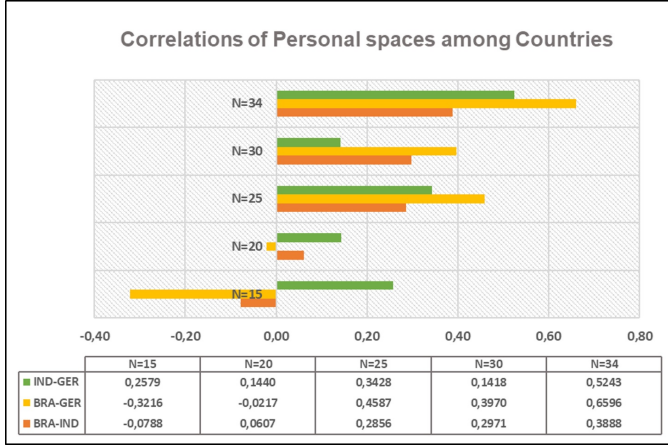


Fig. 8. **Correlations of personal space:** correlations measured for the examined countries.

VII. FINAL REMARKS

We proposed the **Big-Four Geometrical Dimensions** or just **Big4GD**, a model containing a set of pedestrian and crowd features grouped into four dimensions: *I - Physical*, *II - Social*, *III - Personal and Emotional* and *IV - Cultural*. This four dimensions enclose characteristics in different levels, as pedestrian, group and crowd. This model is an approach to detect these dimensions at a specific moment in a certain physical space, from the geometric point of view and not a scientific tool for assessing cultural, personality or emotional profiles of population.

Based on geometric characteristics derived from pedestrian trajectories, we propose a way of characterizing pedestrians and groups of individuals in crowds, allowing the comparison of each other to find differences between one crowd and another. Based on a series of experiments, we were able to validate our model and verify that our approach succeeds in extracting the information from the crowds and their individuals. In this way, regarding personality and cultural aspects, we successfully compared our results with Psychology literature, where several studies aimed to analysis human behavior. It is important to notice that, even if the literature measured these dimensions by considering a different type of information (subjective responses of individuals collected through questionnaires), the results obtained from our approach using geometrical information indicate that our model generates

coherent information when compared to data provided in available literature

VIII. PUBLICATIONS AND DISTINCTIONS

- R. M. Favaretto, S. R. Musse and A. B. Costa. **Emotion, Personality and Cultural Aspects in Crowds: Towards a Geometrical Mind**. Springer Nature. 2019. **Book** [33].
- R. M. Favaretto, R. R. dos Santos, S. R. Musse, F. Vilanova and A. B. **Investigating cultural aspects in the fundamental diagram using convolutional neural networks and virtual agent simulation**. Comput Anim Virtual Worlds. 2019. **Journal** [34].
- R. M. Favaretto, P. Knob, S. R. Musse, F. Vilanova, and A. B. Costa. **Detecting personality and emotion traits in crowds from video sequences**. Machine Vision and Applications, 5:9991012, 2018. **Journal** [27].
- R. M. Favaretto, L. Dihl, R. Barreto and S. R. Musse. **Using group behaviors to detect hofstede cultural dimensions**. In International Conference on Image Processing (ICIP). Phoenix, AZ, 2016. **Conference** [30].
- R. M. Favaretto, L. Dihl, and S. R. Musse. **Detecting crowd features in video sequences**. SIBGRAPI. So Jos dos Campos, SP, 2016. **Conference** [25].
- R. M. Favaretto, L. Dihl, S. R. Musse, F. Vilanova, and A. B. Costa. **Using big-five personality model to detect cultural aspects in crowds**. SIBGRAPI. Niteroi, RJ, 2017. **Conference** [24].
- P. Knob, M. Alcantara, E. Testa, R. Favaretto, G. Lima, L. Dihl, and S. R. Musse. **Generating background NPCs motion and grouping behavior based on real video sequences**. Entertainment Computing, 2018. **Journal** [35].
- V. Araujo, R. M. Favaretto, P. Knob, S. R. Musse, F. Vilanova, and A. B. Costa. **How Much Do You Perceive This?: An Analysis on Perceptions of Geometric Features, Personalities and Emotions in Virtual Humans**. ACM International Conference on Intelligent Virtual Agents. Paris, FR, 2019. **Conference** [36].
- M. Alcantara, E. Testa, G. L. da Silva, R. Favaretto, L. Dihl, and S. Musse. **Generating background population for games based on real video sequences**. SBGames. So Paulo, SP, 2016. SBGames. **Conference** [37].
- L. Dihl, E. Testa, P. Knob, G. Lima, R. M. Favaretto, M. Alcantara, and S. R. Musse. **Generating cultural characters based on hofstede dimensions**. In Virtual Humans and Crowds for Immersive Environments (VHCIE). Los Angeles, CA, 2017. **Conference** [38].
- P. Knob, V. F. de A. Araujo, R. M. Favaretto, and S. R. Musse. **Visualization of interactions in crowd simulation and video sequences**. SBGames. Foz do Iguacu, PR, 2018. **Conference** [39].
- R. M. Favaretto. **Geomind**. A software to extract emotion, personality and cultural aspects from pedestrians in video sequences. 2019. **Software** [40].
- R. M. Favaretto. **Cultural Crowds Dataset**. A dataset with videos from various countries. 2018. **Dataset** [41].

REFERENCES

- [1] J. Jacques Junior, S. R. Musse, and C. Jung, "Crowd analysis using computer vision techniques," *IEEE Signal Processing Magazine*, vol. 27, pp. 66–77, 2010.
- [2] Z. Shakhosini and M. Sarvi, "Pedestrian crowd flows in shared spaces: investigating the impact of geometry based on micro and macro scale measures," *Transportation Research Part B: Methodological*, vol. 112, pp. 57–87, 2019.
- [3] W. Zhao, Z. Zhang, and K. Huang, "Gestalt laws based tracklets analysis for human crowd understanding," *Pattern Recognition*, vol. 75, pp. 112–127, 2018.
- [4] W. Li, Z. Di, and J. M. Allbeck, "Crowd distribution and location preference," *Computer Animation and Virtual Worlds*, vol. 23, pp. 343–351, 2012.
- [5] U. Chattaraj, A. Seyfried, and P. Chakroborty, "Comparison of pedestrian fundamental diagram across cultures," *Advances in Complex Systems*, vol. 12, pp. 393–405, 2009.
- [6] G. Hofstede, *Culture's consequences: comparing values, behaviors, institutions, and organizations across nations*. Thousand Oaks, CA: Sage Publications, 2001.
- [7] P. Costa and R. McCrae, *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI)*. New York, NY: Psychological Assessment Resources, 1992.
- [8] R. R. McCrae and P. L. Costa, *Toward a new generation of personality theories: Theoretical contexts for the five-factor model*. New York, NY: Guilford Press, 1996, ch. 3, pp. 51–87.
- [9] A. Ortony, G. L. Clore, and A. Collins, *The cognitive structure of emotions*. New York, NY: Cambridge university press, 1990.
- [10] A. K. Chandran, A. P. Loh, and P. Vadakkepat, "Identifying social groups in pedestrian crowd videos," in *Eighth International Conference on Advances in Pattern Recognition (ICAPR)*. Kolkata, IN: IEEE, 2015, pp. 1–6.
- [11] F. Solera, S. Calderara, and R. Cucchiara, "Structured learning for detection of social groups in crowd," in *10th IEEE International Conference on Advanced Video and Signal Based Surveillance*. Krakow, PL: IEEE, 2013, pp. 7–12.
- [12] A. Seyfried and A. Schadschneider, "Fundamental diagram and validation of crowd models," in *Proceedings of the 8th International Conference on Cellular Automata for Research and Industry*. Berlin, GE: Springer-Verlag, 2008, pp. 563–566.
- [13] A. Seyfried, M. Boltes, J. Kähler, W. Klingsch, A. Portz, T. Rupperecht, A. Schadschneider, B. Steffen, and A. Winkens, *Enhanced empirical data for the fundamental diagram and the flow through bottlenecks*. Berlin, GE: Springer Berlin Heidelberg, 2010, ch. 11, pp. 145–156.
- [14] S. Narang, A. Best, S. Curtis, and D. Manocha, "Generating pedestrian trajectories consistent with the fundamental diagram based on physiological and psychological factors," *PLOS ONE*, vol. 10, pp. 1–17, 2015.
- [15] A. Best, S. Narang, S. Curtis, and D. Manocha, "Densesense: interactive crowd simulation using density-dependent filters," in *Proceedings of the ACM SIGGRAPH/Eurographics Symposium on Computer Animation*. Switzerland, SW: Eurographics Association, 2014, pp. 97–102.
- [16] D. Helbing, A. Johansson, and H. Z. Al-Abideen, "Dynamics of crowd disasters: an empirical study," *Physical Review E*, vol. 75, pp. 461091–461097, 2007.
- [17] L. Saifi, A. Boubetra, and F. Nouioua, "An approach for emotions and behavior modeling in a crowd in the presence of rare events," *Adaptive Behavior*, vol. 24, pp. 428–445, 2016.
- [18] D. Lala, S. Thovuttikul, and T. Nishida, "Towards a virtual environment for capturing behavior in cultural crowds," in *Sixth International Conference on Digital Information Management*. Melbourne, AU: IEEE, 2011, pp. 310–315.
- [19] N. Fridman, A. Zilka, and G. A. Kaminka, "The impact of cultural differences on crowd dynamics in pedestrian and evacuation domains," Bar Ilan University, (Technical Report), 2011.
- [20] S. Cao, L. Lian, M. Chen, M. Yao, W. Song, and Z. Fang, "Investigation of difference of fundamental diagrams in pedestrian flow," *Physica A: Statistical Mechanics and its Applications*, vol. 506, pp. 661–670, 2018.
- [21] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*. Kauai, HI: IEEE Computer Society, 2001, pp. 511–518.
- [22] J. Bins, L. L. Dohl, and C. R. Jung, "Target tracking using multiple patches and weighted vector median filters," *Journal of Mathematical Imaging and Vision*, vol. 45, pp. 293–307, 2013.
- [23] E. T. Hall, *The Hidden Dimension*. Garden City, NY: Doubleday, 1966.
- [24] R. M. Favaretto, L. Dohl, S. R. Musse, F. Vilanova, and A. B. Costa, "Using big-five personality model to detect cultural aspects in crowds," in *Conference on Graphics, Patterns and Images (SIBGRAPI)*. Niteroi, RJ: IEEE, 2017, pp. 223–229.
- [25] R. M. Favaretto, L. Dohl, and S. R. Musse, "Detecting crowd features in video sequences," in *Proceedings of Conference on Graphics, Patterns and Images (SIBGRAPI)*. So Jos dos Campos, SP: IEEE Computer Society, 2016, pp. 201–208.
- [26] P. Costa and R. R. McCrae, *NEO PI-R - Inventrio de Personalidade NEO Revisado*. So Paulo, SP: Vetor, 2007.
- [27] R. M. Favaretto, P. Knob, S. R. Musse, F. Vilanova, and A. B. Costa, "Detecting personality and emotion traits in crowds from video sequences," *Machine Vision and Applications*, vol. 30, no. 5, pp. 999–1012, Jul 2019.
- [28] V. Araujo, R. Migon Favaretto, P. Knob, S. Raupp Musse, F. Vilanova, and A. Brandelli Costa, "How much do you perceive this?: An analysis on perceptions of geometric features, personalities and emotions in virtual humans," in *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*. New York, NY, USA: ACM, 2019, pp. 179–181.
- [29] G. Hofstede, G. J. Hofstede, and M. Minkov, *Cultures and organizations: Software of the mind*. New York, NY: McGraw-Hill, 1991.
- [30] R. M. Favaretto, L. Dohl, R. Barreto, and S. R. Musse, "Using group behaviors to detect hofstede cultural dimensions," in *International Conference on Image Processing (ICIP)*. Phoenix, AZ: IEEE, 2016, pp. 2936–2940.
- [31] F. Vilanova, F. M. Beria., A. B. Costa, and S. H. Koller, "Deindividuation: From le bon to the social identity model of deindividuation effects," *Cogent Psychology*, vol. 4, pp. 1–21, 2017.
- [32] G. L. Bon, *A Psicologia das Multides*. Paris, FR: Presses Universitaires de France, 1986.
- [33] S. R. M. Rodolfo M. Favaretto and A. B. Costa, *Emotion, Personality and Cultural Aspects in Crowds: Towards a Geometrical Mind*. London, UK: Springer Nature, 2019.
- [34] R. Migon Favaretto, R. Rosa dos Santos, S. Raupp Musse, F. Vilanova, and A. Brandelli Costa, "Investigating cultural aspects in the fundamental diagram using convolutional neural networks and virtual agent simulation," *Computer Animation and Virtual Worlds*, vol. 30, no. 3–4, p. e1899, 2019, e1899 cav.1899.
- [35] P. Knob, M. Alcantara, E. Testa, R. Favaretto, G. Lima, L. Dohl, and S. R. Musse, "Generating background npcs motion and grouping behavior based on real video sequences," *Entertainment Computing*, vol. 27, pp. 179–187, 2018.
- [36] V. Araujo, R. Migon Favaretto, P. Knob, S. Raupp Musse, F. Vilanova, and A. Brandelli Costa, "How much do you perceive this?: An analysis on perceptions of geometric features, personalities and emotions in virtual humans," in *Proceedings of the 19th ACM International Conference on Intelligent Virtual Agents*, ser. IVA '19. New York, NY, USA: ACM, 2019, pp. 179–181. [Online]. Available: <http://doi.acm.org/10.1145/3308532.3329454>
- [37] M. Alcantara, E. Testa, G. L. da Silva, R. Favaretto, L. Dohl, and S. Musse, "Generating background population for games based on real video sequences," in *Brazilian Symposium on Games and Digital Entertainment (SBGames)*. So Paulo, SP: SBGames, 2016, pp. 66–72.
- [38] L. Dohl, E. Testa, P. Knob, G. Lima, R. M. Favaretto, M. Alcantara, and S. R. Musse, "Generating cultural characters based on hofstede dimensions," in *Virtual Humans and Crowds for Immersive Environments (VHCIE)*. Los Angeles, CA: IEEE, 2017, pp. 1–6.
- [39] P. Knob, V. F. de A. Araujo, R. M. Favaretto, and S. R. Musse, "Visualization of interactions in crowd simulation and video sequences," in *XVII Brazilian Symposium on Games and Digital Entertainment (SBGames)*. Foz do Iguau, PR: SBGames, 2018, pp. 601–610.
- [40] R. M. Favaretto, "Geomind: a software to extract emotion, personality and cultural aspects from video sequences." 2019. [Online]. Available: <https://www.rmfavaretto.pro.br/geomind>
- [41] —, "Cultural crowds dataset: a set of videos from several countries," 2018. [Online]. Available: <https://www.rmfavaretto.pro.br/vhlab/datasets.php>