Aggressive, Tense, or Shy?
Identifying Personality Traits from Crowd Videos

Aniket Bera, Tanmay Randhavane, and Dinesh Manocha

Department of Computer Science, University of North Carolina at Chapel Hill

ab@cs.unc.edu

Abstract

We present a real-time algorithm to automatically classify the dynamic behavior or personality of a pedestrian based on his or her movements in a crowd video. Our classification criterion is based on Personality Trait Theory. We present a statistical scheme that dynamically learns the behavior of every pedestrian in a scene and computes that pedestrian's motion model. This model is combined with global crowd characteristics to compute the movement patterns and motion dynamics, which can also be used to predict the crowd movement and behavior. We highlight its performance in identifying the personalities of different pedestrians in low- and high-density crowd videos. We also evaluate the accuracy by comparing the results with a user study.

1 Introduction

Modeling and classifying the behavior of different pedestrians in a crowd is an important problem in various domains including psychology, robotics, pedestrian dynamics, and behavior learning. Even simple tasks like walking towards a goal position involve several complex decisions such as figuring out the most efficient path or route, and choosing between the various available paths to avoid collisions. According to *Convergence Theory* [Turner and Killian, 1987], a well-known approach used in sociology and economics, crowd behavior is not a sole product of the crowd itself; rather, it is defined by the individual pedestrians in that crowd. As a result, it is important to accurately predict the behavior of individuals and their interactions with the environment to capture realistic, heterogeneous crowd behaviors.

Recent advances in sensor technologies have made it easier to capture high resolution videos of pedestrians and crowds. Moreover, surveillance cameras are frequently used in public places and buildings for monitoring human behaviors. In this paper, we address the problem of classifying the behaviors of different pedestrians in a crowd video based on their movement patterns and use these patterns for crowd behavior prediction. Besides surveillance, these techniques are also useful for architectural design and collision-free navigation of robots or autonomous vehicles in crowded scenarios.

Figure 1: Crowd Behavior Learning/Prediction: Our approach can automatically classify the behavior of each pedestrian in a large crowd. We highlight its application for the 2017 Presidential Inauguration crowd video at the National Mall at Washington, DC (courtesy PBS): (1) original aerial video footage of the dense crowd; (2) a synthetic rendering of pedestrians in the red square based on their personality classification: aggressive (**orange**), shy (**black**), active (**blue**), tense (**purple**), etc; (3) a predicted simulation of 1M pedestrians in the space with a similar distribution of personality traits.

Many factors including biological, developmental, and situational variations, along with individual personalities, govern people's overall behavior. We mainly focus on capturing the variations in behavior that arise as humans navigate the physical world and avoid collisions. In general, categorizing the variety of personalities that humans exhibit can be hard. Psychologists have proposed different models to represent these variations, but they have some limitations [Harvey *et al.*, 1995]. Therefore, we base our classification on *Person-*

ality Trait Theory, which proposes that a wide range of variations in behavior is primarily the result of a small number of underlying traits. It is also important to model many external or environmental factors, including surrounding pedestrians and the crowd's movement flow for estimation and prediction.

Main Results: We present a novel learning algorithm to classify pedestrian behaviors based on their movement patterns. We extract the trajectory of each pedestrian in a video and use a combination of Bayesian learning and pedestrian dynamics techniques to compute the local and global characteristics at interactive rates. The local characteristics include the time-varying motion model that is used to compute the personality traits. We also present new statistical algorithms to learn high-level characteristics and global movement patterns. We combine these characteristics with Eysenck's 3-factor PEN model [Eysenck and Eysenck, 1985] and characterize the personality into six weighted behavior classes: *aggressive, assertive, shy, active, tense,* and *impulsive*. We also use the individual personalities to predict the state of the crowd under different environmental scenarios.

To the best of our knowledge, this is the first approach that can automatically identify the behavior of each pedestrian in a crowd. We have evaluated its accuracy with a user study (88.48%) and evaluated its performance on different videos with tens of pedestrians. One example is the large crowd gathered in Washington, DC for the Presidential Inauguration (January 2017) using *PBS HD* video footage (see Figure 1). Our approach offers many benefits:

1. Robust: Our approach is robust, can account for noise in the pedestrian trajectories, and classifies the behavior using time-varying pedestrian movement dynamics.

2. General: Our approach is applicable to indoor and outdoor crowd videos and makes no assumption about their size or density.

3. Crowd Analysis and Prediction: Our approach can be used to analyze and estimate the future movement or behavior of the crowd. Furthermore, it can be used to predict different scenarios based on the behaviors and global characteristics, e.g., the distribution and density of a large crowd at the National Mall in Figure 1.

The rest of the paper is organized as follows. Section 2 provides an overview of related work in video-based crowd analysis and personality models. We introduce the terminology and present our algorithm for computing the local and global characteristics in Section 3. We highlight the performance on challenging benchmarks and describe results from our user evaluation in Section 4.

2 Related Work

In this section, we give a brief overview of prior work on video-based crowd analysis, behavior classification, and personality models.

2.1 Video-Based Crowd Analysis

There is extensive work in computer vision, multimedia, and robotics that analyzes the behaviors and movement patterns in crowd videos, as surveyed in [Li *et al.*, 2015; Borges
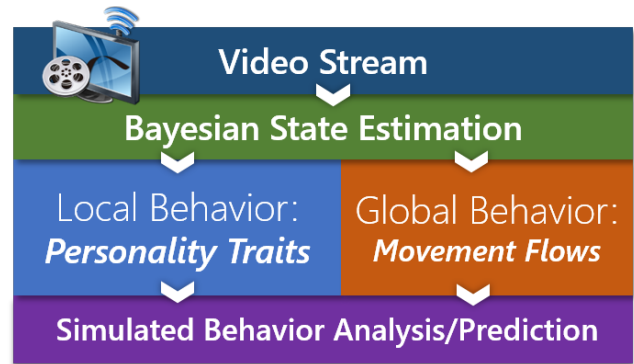

Figure 2: Our method takes a streaming crowd video as an input. We compute the state of pedestrians in the crowd, as explained in Section 3. Based on the state information, we learn local and global behavior properties, which are combined for behavior classification and prediction.

et al., 2013]. The main objectives of these work include human behavior understanding and crowd activity recognition for detecting abnormal behaviors [Hu *et al.*, 2004; Bera *et al.*, 2016]. Many of these methods use a large number of training videos to learn the patterns offline [Zen and Ricci, 2011; Solmaz *et al.*, 2012]. Other methods utilize motion models to learn crowd behaviors [Pellegrini *et al.*, 2012] or use machine learning algorithms [Zhou *et al.*, 2012; Cheung *et al.*, 2016]. Some techniques focus on classifying the most common behavior patterns in a given scene using offline learning. These include activity prototypes using a convex learning algorithm [Zen and Ricci, 2011] and detection of popular behavior patterns like bottlenecks, fountainheads, lanes, arches, and blocking [Solmaz *et al.*, 2012].

Crowd behavior learning using motion or simulation has been used for different applications. Parameter learning has been used to predict pedestrian motion for tracking [Pellegrini *et al.*, 2012]. However, these techniques use either manual selection or offline learning techniques to estimate the goal positions. Other researchers have used low-density tracking data to learn agent intentions [Musse *et al.*, 2007] or use online Bayesian motion-prediction methods for human-robot interactions, data-driven crowd simulation [Kim *et al.*, 2016], and offline training [Zhong *et al.*, 2015].

2.2 Pedestrian Behavior Modeling

Different approaches have been used to model pedestrian behavior. [Funge *et al.*, 1999] use cognitive modeling to empower agents to plan and perform high-level tasks [Godoy *et al.*, 2016]. Other approaches use personality models to simulate the behavior of pedestrians and crowds including the OCEAN model [Durupinar *et al.*, 2011], the MBTI model [Salvit and Sklar, 2011], and Personality Trait theory [Guy *et al.*, 2011] and General Adaptation Syndrome Theory [Kim *et al.*, 2012]. However, these methods only take into account local motion models, not the global characteristics.

3 Pedestrian Behavior Computation

In this section, we present our interactive algorithm, which classifies pedestrian behavior using 2D, real-world trajectories that are extracted from video. Our approach can be combined with almost any real-time pedestrian tracker that works on dense crowd videos. Figure 2 gives an overview of our approach. Our method takes a live or streaming crowd video as an input. We extract the initial set of pedestrian trajectories using an online pedestrian tracker. We learn the pedestrian motion model parameters using statistical methods and learn global and local behavior characteristics. These can be used to predict the future state of the pedestrian or the overall crowd.

3.1 Symbols and Notation

We introduce the terminology and symbols used in the rest of the paper. We use the term *pedestrian* to refer to an independent individual or agent in the crowd. We use the notion of *state* to specify the trajectory and behavior characteristics of each pedestrian; these characteristics also govern the positional movements of a pedestrian on a 2D plane. We use the symbol $\mathbf{x} \in \mathbb{R}^5$ to refer to a pedestrian's state:

$$\mathbf{x} = [\mathbf{p} \ \mathbf{v}^c \ \mathbf{v}^{pref}]^\mathbf{T}, \tag{1}$$

where $\mathbf{p}$ is the pedestrian's position, $\mathbf{v}^c$ is its current velocity, and $\mathbf{v}^{pref}$ is the *preferred velocity* on a 2D plane. The preferred velocity is the optimal velocity that a pedestrian would take to achieve its intermediate goal if there were no other pedestrians or obstacles in the scene. In practice, $\mathbf{v}^{pref}$ tends to be different from $\mathbf{v}^c$ for a given pedestrian. We use the symbol $\mathbf{S}$ to denote the current state of the environment, which corresponds to the states of all the other pedestrians and the current positions of the obstacles in the scene. The state of the crowd, which consists of individual pedestrians, is a union of the set of each pedestrian's state $\mathbf{X} = \bigcup_i \mathbf{x_i}$, where subscript i denotes the ith pedestrian. The difference between $\mathbf{v}^{pref}$ and $\mathbf{v}^c$ provides partial information about the local interactions between a pedestrian and the rest of the environment

Motion Model: $\mathbf{P} \in \mathbb{R}^5$ denotes the set of parameters for the motion model. The motion model corresponds to the local navigation rule or scheme that each pedestrian uses to avoid collisions with other pedestrians or obstacles. Our formulation is based on the RVO velocity-based motion model [Van den Berg *et al.*, 2008]. In this model, the motion of each pedestrian is governed by these five characteristics: *Neighbor Dist, Maximum Neighbors, Planning Horizon, Radius, and Preferred Speed*. Our approach can also be combined with other models based on social forces or Boids.

3.2 Personality Trait Classification

Psychologists have proposed various ways of characterizing the personalities exhibited by pedestrians. Our work builds on Trait Theories of Personalities, a theory that categorizes people's behavior based on a small number of personality traits [Pervin, 2003]. The overall goal is to automatically classify every pedestrian in a crowd. In particular, we characterize each pedestrian behavior based on a weighted combination of
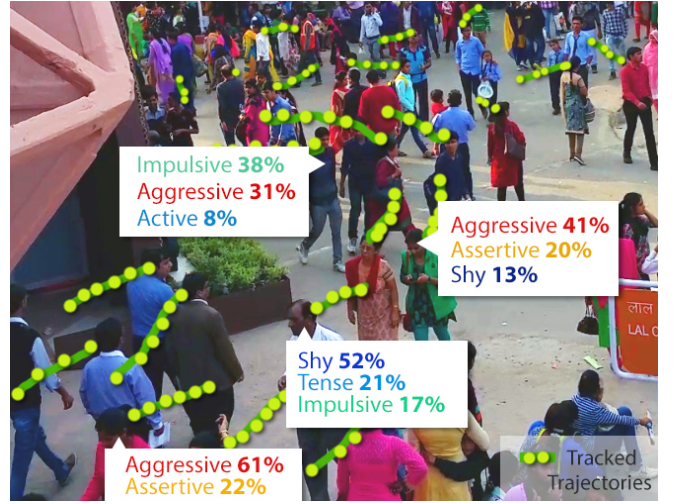

Figure 3: Personality Classification: We identify the personality of each tracked pedestrian based on pedestrian dynamics and the motion model. Each pedestrian is automatically classified using a weighted combination of different personality traits.

different personality traits that are inferred based on his or her movement pattern and interactions with other pedestrians and obstacles in the environment. We use the well-known Personality Trait Theory from psychology and the Eysenck 3-factor model [Eysenck and Eysenck, 1985] to classify such behaviors. This model identifies three major factors that characterize the personality: *Psychoticism*, *Extraversion*, and *Neuroticism* (commonly referred to as **PEN**). Each of these three traits has been linked to a biological basis such as the levels of testosterone, serotonin, and dopamine present in ones body.

In our case, each pedestrian's personality is identified based on how they exhibit each of these three traits. Each pedestrian is then classified into one of six weighted behavior classes: aggressive, assertive, shy, active, tense, and impulsive (see Fig. 3). We chose these six particular behavior characteristics because they are useful in describing pedestrians' behaviors, and span the space covered by the PEN model, with at least two adjectives for each PEN trait [Pervin, 2003]. We classify each pedestrian's behavior using these traits, and our fundamental assumption is that these behaviors are captured by a weighted combination of these six traits.

A key issue in this formulation is defining a mapping between the five motion model parameters for the RVO model described in Section 3.1 and these six traits. We make use of the data-driven mapping presented in [Guy *et al.*, 2011] to derive such a mapping that adopts the results of a user study and derives a linear model of the mapping as:

$$\mathbf{B} = \begin{pmatrix} Aggressive \\ Assertive \\ Shy \\ Active \\ Tense \\ Impulsive \end{pmatrix} = \mathbf{M}_{mat} * \begin{pmatrix} \frac{1}{13.5}(Neighbor\ Dist - 15) \\ \frac{1}{49.5}(Max.\ Neighbors - 10) \\ \frac{1}{14.5}(Planning\ Horiz. - 30) \\ \frac{1}{0.85}(Radius - 0.8) \\ \frac{1}{0.5}(Pref.\ Speed - 1.4) \end{pmatrix}$$

where,

$$\mathbf{M_{mat}} = \begin{pmatrix} -0.02 & 0.32 & 0.13 & -0.41 & 1.02 \\ 0.03 & 0.22 & 0.11 & -0.28 & 1.05 \\ -0.04 & -0.08 & 0.02 & 0.58 & -0.88 \\ -0.06 & 0.04 & 0.04 & -0.16 & 1.07 \\ 0.10 & 0.07 & -0.08 & 0.19 & 0.15 \\ 0.03 & -0.15 & 0.03 & -0.23 & 0.23 \end{pmatrix}.$$

Even though our approach is general, the mapping (M_{mat}) is specific to the RVO motion model and the user study described in [Guy *et al.*, 2011].

3.3 Global Characteristics

The trajectories extracted from a real-world video tend to be noisy and may have incomplete tracks [Enzweiler and Gavrila, 2009]. Therefore we use the Bayesian-inference technique to compensate for any errors and to compute the state of each pedestrian. We use an Ensemble Kalman Filter (EnKF) and Expectation Maximization (EM) to estimate the most likely state $\mathbf{x}$ of each pedestrian.

Our approach extends the method presented in [Kim *et al.*, 2016]. The global dynamics consist of factors that govern pedestrians' trajectory behaviors in a group or crowd, i.e., the factors that influence a pedestrian's overall movement or flow. We use two main components to describe the global dynamics: start point of each pedestrian in the scenario and movement flows. We then use them to analyze the global behavior of each pedestrian. Formally, we represent these dynamic characteristics for each pedestrian with a vector-valued function, $f()$, whose initial value is determined by the function, $E()$:

$$\mathbf{x}^{t+1} = f(t, \mathbf{x}^t) = [P(\mathbf{x}^t)\, I(\mathbf{x}^t)\, G(t, \mathbf{x}^t)]^{\mathbf{T}}; \quad \mathbf{x}^0 = E(t^0).$$

For each pedestrian in the crowd, the function $G : \mathbb{R} \times \mathbb{R}^6 \times \mathbb{S} \to \mathbb{R}^2$ maps time t, the current state of the pedestrian $\mathbf{x} \in \mathbf{X}$, and the current state of the environment $\mathbf{S} \in \mathbb{S}$ to a preferred velocity $\mathbf{v}^{pref}$. Function $I : \mathbb{R}^6 \times \mathbb{S} \to \mathbb{R}^2$ represents the RVO motion model that is used to compute the current velocity $\mathbf{v}^c$ for collision-free interactions with other pedestrians and obstacles. The function $P : \mathbb{R}^2 \to \mathbb{R}^2$ computes the position given $\mathbf{v}^c$ and $E : \mathbb{R} \to \mathbb{R}^2$ computes the initial position for time t_0, which is the time at which a particular pedestrian enters the environment. The three components of the pedestrian dynamics (start point, movement flow, and local collision-free navigation) are mapped to the functions $E()$, $G()$, and $I()$, respectively. We use Bayesian inference to compute $E()$ and $G()$ from the 2D trajectory data of the pedestrians.

Movement Feature The movement features describe the characteristics of the trajectory behavior at a certain position at time frame t. These movement features of different pedestrians are grouped together and form a cluster that represents the movement flow. The characteristics include the movement of the pedestrians during the past w frames, which we call the *time window*, and the intended direction of the movement (i.e. the preferred velocity) at this position. In our case, the movement feature vector is represented as a six-dimensional vector:

$$\mathbf{g} = [\mathbf{p}\ \mathbf{v}^{avg}\ \mathbf{v}^{pref}]^T, \tag{2}$$

where $\mathbf{p}$, $\mathbf{v}^{avg}$, and $\mathbf{v}^{pref}$ are each two-dimensional vectors that correspond to the current position, average velocity during past w frames, and estimated preferred velocity computed as part of state estimation, respectively. $\mathbf{v}^{avg}$ can be computed from $(\mathbf{p}^t - \mathbf{p}^{t-w})/w * dt$, where dt is the time-step.

Movement Flow Clustering At every w steps, we compute the new behavior features for each pedestrian using Equation 2. We group similar features and find K most common behavior patterns, which correspond to the *movement flow clusters*. We use recently observed behavior features to learn the time-varying movement flow. We use the k-means data clustering algorithm to classify these features into K movement flow clusters. In our case, K and N^f are user defined values that represent the total number of the clusters and the total number of collected behavior features, respectively, and $K \leq N^f$. A set of movement-flow clusters $G = \{G_1, G_2, ..., G_K\}$ is computed as follows:

$$\underset{G}{\mathrm{argmin}} \sum_{k=1}^{K} \sum_{g_i \in G_k} dist(b_i, \mu_k), \tag{3}$$

where g_i is a movement feature vector, μ_k is a centroid of each flow cluster, and $dist(g_i, \mu_k)$ is a distance measure between the arguments. In our case, the distance between two feature vectors is computed as

$$\begin{aligned} dist(g_i, g_j) &= c_1 \left\| \mathbf{p}_i - \mathbf{p}_j \right\| \\ &+ c_2 \left\| (\mathbf{p}_i - \mathbf{v}_i^{avg} w\, dt) - (\mathbf{p}_j - \mathbf{v}_j^{avg} w\, dt) \right\| \\ &+ c_3 \left\| (\mathbf{p}_i + \mathbf{v}_i^{pref} w\, dt) - (\mathbf{p}_j - \mathbf{v}_j^{pref} w\, dt) \right\|, \end{aligned}$$

which corresponds to the weighted sum of the distance among three points: current positions, previous positions, and estimated future positions.

Estimation of Start Points Start points for each point correspond to the estimated position when that pedestrian enters the scene. These starting positions and timings for each position are very important and are used to compute the global behavior. We use a multivariate Gaussian mixture model to learn the time-varying distribution of the start points. We define $E()$ as the function that provides a position sampled from the learned distributions. We assume that the distribution of start points, e, from which the function $E()$ samples, is a mixture of J components and that each of the components is a multivariate Gaussian distribution of a two-dimensional random variable, $\mathbf{p}$, with a set of parameters $\Theta = (\alpha_1, \cdots, \alpha_J, \theta_1, \cdots, \theta_J)$:

$$e(\mathbf{p}|\Theta) = \sum_{j=1}^{J} \alpha_j e_j(\mathbf{p}|\mu_j, \theta_j), \tag{4}$$

$$e_j(\mathbf{p}; \theta_j) = \frac{1}{2\pi |\Sigma_j|^{1/2}} exp(-\frac{1}{2}(\mathbf{p} - \mu_j)^T \Sigma_j^{-1} (\mathbf{p} - \mu_j)). \tag{5}$$

Each component e_j is a Gaussian distribution given by the parameters $\theta_j = (\mu_j, \Sigma_j)$, where μ_j is the mean of the component j and Σ_j is a 2×2 covariance matrix. α_j is a mixture weight, which is the probability that a point $\mathbf{p}$ belongs to the component j. $\alpha_j \in [0, 1]$ for all i and the sum of α_js are constrained to 1 ($1 = \sum_{j=1}^{J} \alpha_j$). From an initial guess of

the parameters θ_j, we perform EM to learn these parameters $\theta_j = (\mu_j, \Sigma_j)$ from the given start points collected from the real pedestrian trajectories.

3.4 Simulated Crowd Behavior Analysis & Prediction

Given a video, we can classify the personality trait, B, of each pedestrian (PTC) and compute the global features (GMD) corresponding to the start points, e, and the movement flows, G. We can combine this local and global information to predict the future position or state of each pedestrian and learn the shape, distribution, or behavior of the crowd in different environmental conditions. These different conditions may correspond to a change in the obstacle locations, an increase or decrease in the number of pedestrians in the scene, or a change in crowd density, but still maintain the original distribution of pedestrian behaviors and movement flows. We demonstrate our approach's performance on the *PBS* video stream from the 2017 Presidential Inauguration ceremony at Washington, DC, USA. We extract a representative sample of the crowd by selecting 130 pedestrians from a camera angle and learn their behaviors. As part of our crowd prediction, we changed the number of pedestrians in the scenario (e.g., 1 million pedestrians), and estimated the distribution and shape of the resulting crowd at the National Mall, as shown in Figure 1. The resulting crowd of 1M pedestrians has the same behavior classification as the original 130 representative pedestrians.

4 Performance and User Evaluation

In this section, we highlight the performance of our algorithms on different crowd videos. Furthermore, we evaluate the accuracy of our personality classification algorithm with a user study.

4.1 Performance Evaluation

We have applied our novel algorithms to the 2D pedestrian trajectories generated and extracted from different crowd videos, as shown in Table 1. The pedestrian density in these crowd videos varies from low-density (less than 1 pedestrian per square meter) to medium-density (1-2 pedestrians per square meter), to high-density (more than 2 pedestrians per square meter). We have implemented our system on a Windows 10 desktop PC with Intel Xeon E5-1620 v3 with 16 GB of memory and we use four cores for PTC and GMD computations.

4.2 User Evaluation

To evaluate the performance of our personality classification, we compare the results of PTC with a user study on the same set of videos and pedestrians. We present the details and highlight the results.

Study Goals and Expectations The aim of the study was to compare the results of our personality trait computation algorithm (PTC) to the perception of personality by human users for the given videos.

Experimental Design The participants of our web-based user study were recruited from students and staff in a university. Each participant was shown 10 different videos of pedestrians walking among crowds and asked to label them based

Scenario	Analysed Pedestrians	Input Frames	Avg. time PTC	Avg. time GMD
Manko	42	373	0.034	3.72e-01
Marathon	18	450	0.041	0.98E-04
Explosion	19	238	0.033	3.12E-06
Street	147	9014	0.022	4.13E-08
TrainStation	200	999	0.061	5.11E-08
ATC Mall	50	7199	0.037	2.28E-01
IITF-1	18	450	0.041	3.11E-03
IITF-3	19	238	0.046	2.74E-04
IITF-5	18	450	0.056	2.98E-03
NPLC-1	19	238	0.039	1.31E-04
NPLC-3	18	450	0.031	4.16E-03
NDLS-2	19	238	0.049	3.00E-04
2017 Presidential Inauguration	130*	1927	0.87	0.38

Table 1: Performance of PTC (Personality Trait computation) and GMD (Global Movement Dynamics) algorithms on different crowd videos. We highlight the number of pedestrians used for personality classification, the number of video frames used for extracted trajectories, and the running time (in seconds). In the PBS Presidential Inauguration video, we chose around 130 representative pedestrians in the video for analysis and prediction.

on personality adjectives. At the beginning of each video, a specific pedestrian was highlighted by a yellow circle and, over the next few seconds, that pedestrian would navigate the crowded scene. Participants were given written instructions to carefully observe the highlighted pedestrian's trajectory and its interactions with the other pedestrians. They were advised not to take into account other factors corresponding to facial expressions, visual appearance, etc. After observing the pedestrian trajectory in each video, participants reported the most appropriate personality trait from the given set: *Aggressive, Shy, Active, Tense, Impulsive, and Assertive*.

Results and Discussion Figure 4 shows the responses of 31 participants for 10 videos. Even though we model an individual's personality as a combination of six personality traits (6-D), evaluating a nominal variable with such a high number of categories is difficult. From the data, we observed that combining two personality traits to reduce the six factor model to a three factor model increases the agreement between participant responses. This three factor model is the **PEN** model (3-D) (explained in Section 3.3) and previous studies have shown that it can also offer sufficiently rich dimensions to characterize personality traits in crowd navigation. We therefore map the six personality traits to the 3-factor PEN model (Figure 5) using a linear mapping as described below.

Trait	Adjectives
Psychoticism	Aggressive, Impulsive
Extraversion	Assertive, Active
Neuroticism	Shy, Tense

Table 2: Correspondence between six personality traits and the PEN model [Pervin, 2003].

Some of the videos still show disagreement between the participants; we attribute this to the inherent features of the
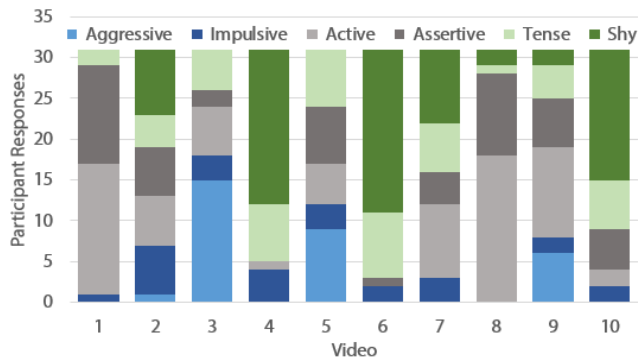
Figure 4: Participant Responses Using the Six Factor Personality Model: Participants were shown 10 different videos of pedestrians walking among crowds. In each video, a single pedestrian was highlighted and participants were asked to report the most appropriate personality trait for that pedestrian. 31 participants reported the most appropriate personality trait from the given set: *Aggressive, Impulsive, Active, Assertive, Tense, and Shy.*
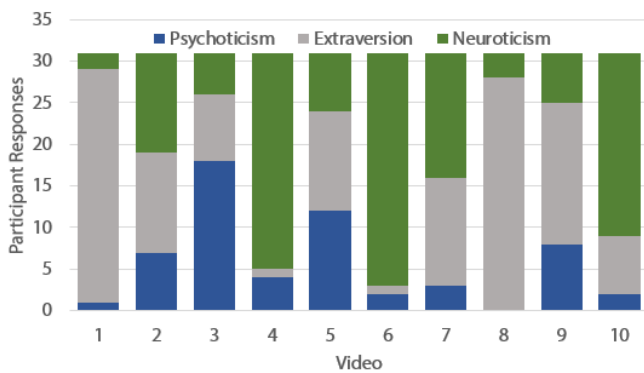

Figure 5: Participant Responses Using the PEN Model: We converted the participant responses to the three factor PEN model to reduce the disagreement. Participant responses were converted to three PEN factors (*Psychoticism, Extraversion, and Neuroticism*) using Table 2.

pedestrian in the videos and do not take these videos (Videos 2, 5, and 7) into account for further analysis.

Figure 6 shows the comparison of participants' responses to the personality traits predicted by our PTC algorithm. Over the 7 crowd videos, we observed an overall accuracy of 76.96% for the most dominant personality trait given by 31 participants. The accuracy increased to 88.48% if we also included the second most dominant personality trait. We also computed a statistical measure, Fleiss' kappa (κ), to assess the reliability of agreement between the participants. A value of $\kappa = 0.4578$ indicated a moderate agreement between the participants' responses based on [Landis and Koch, 1977]. Error analysis of κ revealed that the observed agreement was not accidental ($Z = 35.6444, p < 0.0001$).

5 Conclusions, Limitations, and Future Work

We present a novel algorithm to classify the personalities of pedestrians in a crowd video. We compute the time-varying
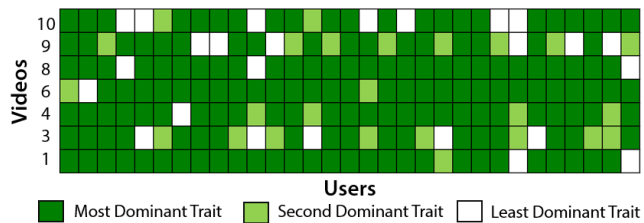

Figure 6: Accuracy of our PTC Algorithm: Our PTC algorithm predicted the most dominant trait for each of the 7 videos. In 76.96% of the cases, participants also chose the most dominant trait as denoted by the dark green color. If we also add the second most dominant trait (denoted by light green), the accuracy increased to 88.48%. This accuracy indicates that our PTC algorithm was able to correctly identify the personality traits as perceived by human participants.

motion model of each pedestrian using Bayesian inference and combine it with Personality Trait Theory. We also compute the global movement features and use them to analyze and simulate the crowd movements or distributions. We evaluate the accuracy with a user study and our results are promising. To the best of our knowledge, this is the first approach for automatic pedestrian personality classification based on their movements in a video.

Our approach has some limitations. The behavior classification is based on personality models and PEN, and may not be sufficient to capture all observed behaviors. We assume that it is possible to extract the trajectory of every pedestrian in a crowd. The global prediction algorithm assumes that the relative distribution of pedestrian behaviors is about the same.

As part of future work, we would like to overcome these limitations and extend our algorithm for anomaly detection and surveillance applications. We would further like to adopt the same data-driven techniques to build mappings from simulation parameters to other personality trait theories, such as the OCEAN model. We would also like to investigate the extent to which our proposed model is appropriate for different cultures. Finally, we would like to use it for autonomous navigation of robots among crowded scenes.

Acknowledgements

References

[Bera *et al.*, 2016] Aniket Bera, Sujeong Kim, and Dinesh Manocha. Interactive crowd-behavior learning for surveillance and training. *IEEE Computer Graphics and Applications*, 36(6):37–45, 2016.

[Borges *et al.*, 2013] P.V.K. Borges, N. Conci, and A. Cavallaro. Video-based human behavior understanding: A survey. *Circuits and Systems for Video Technology, IEEE Transactions on*, 23(11):1993–2008, Nov 2013.

[Cheung *et al.*, 2016] Ernest Cheung, Tsan Kwong Wong, Aniket Bera, Xiaogang Wang, and Dinesh Manocha. Lcrowdv: Generating labeled videos for simulation-based

crowd behavior learning. In *European Conference on Computer Vision*, pages 709–727. Springer, 2016.

[Durupinar *et al.*, 2011] F. Durupinar, N. Pelechano, J.M. Allbeck, U. Gü anddü andkbay, and N.I. Badler. How the ocean personality model affects the perception of crowds. *Computer Graphics and Applications, IEEE*, 31(3):22 – 31, may-june 2011.

[Enzweiler and Gavrila, 2009] Markus Enzweiler and Dariu M Gavrila. Monocular pedestrian detection: Survey and experiments. *IEEE PAMI*, 2009.

[Eysenck and Eysenck, 1985] H.J. Eysenck and M.W. Eysenck. *Personality and individual differences: A natural science approach*. Plenum Press New York, 1985.

[Funge *et al.*, 1999] J. Funge, X. Tu, and D. Terzopoulos. Cognitive modeling: knowledge, reasoning and planning for intelligent characters. In *SIGGRAPH*, pages 29–38. ACM Press, 1999.

[Godoy *et al.*, 2016] Julio Godoy, Ioannis Karamouzas, Stephen J Guy, and Maria Gini. Moving in a crowd: Safe and efficient navigation among heterogeneous agents. In *Proc. Int. Joint Conf. on Artificial Intelligence*, 2016.

[Guy *et al.*, 2011] Stephen J. Guy, Sujeong Kim, Ming C. Lin, and Dinesh Manocha. Simulating heterogeneous crowd behaviors using personality trait theory. In *Symposium on Computer Animation*, pages 43–52. ACM, 2011.

[Harvey *et al.*, 1995] Robert J. Harvey, William D. Murry, and Dean T. Stamoulis. Unresolved issues in the dimensionality of the myers-briggs type indicator. *Educational and Psych. Measurement*, 1995.

[Hu *et al.*, 2004] Weiming Hu, Tieniu Tan, Liang Wang, and Steve Maybank. A survey on visual surveillance of object motion and behaviors. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, 34(3):334–352, 2004.

[Kim *et al.*, 2012] Sujeong Kim, Stephen J Guy, Dinesh Manocha, and Ming C Lin. Interactive simulation of dynamic crowd behaviors using general adaptation syndrome theory. In *Proceedings of the ACM SIGGRAPH Symposium on Interactive 3D Graphics and Games*, 2012.

[Kim *et al.*, 2016] Sujeong Kim, Aniket Bera, Andrew Best, Rohan Chabra, and Dinesh Manocha. Interactive and adaptive data-driven crowd simulation. In *Virtual Reality (VR)*, pages 29–38. IEEE, 2016.

[Landis and Koch, 1977] J Richard Landis and Gary G Koch. The measurement of observer agreement for categorical data. *biometrics*, pages 159–174, 1977.

[Li *et al.*, 2015] Teng Li, Huan Chang, Meng Wang, Bingbing Ni, Richang Hong, and Shuicheng Yan. Crowded scene analysis: A survey. *Circuits and Systems for Video Technology, IEEE Transactions on*, 25(3):367–386, March 2015.

[Musse *et al.*, 2007] Soraia R. Musse, Cludio R. Jung, Julio C. S. Jacques, and Adriana Braun. Using computer vision to simulate the motion of virtual agents. *Computer Animation and Virtual Worlds*, 18(2):83–93, 2007.

[Pellegrini *et al.*, 2012] Stefano Pellegrini, Jrgen Gall, Leonid Sigal, and Luc Gool. Destination flow for crowd simulation. In *Computer Vision ECCV 2012. Workshops and Demonstrations*, volume 7585, pages 162–171. 2012.

[Pervin, 2003] Lawrence Pervin. *The Science of Personality*. Oxford University Press, Oxford, 2003.

[Salvit and Sklar, 2011] J. Salvit and E. Sklar. Toward a Myers-Briggs Type Indicator Model of Agent Behavior in Multiagent Teams. *Multi-Agent-Based Simulation XI*, pages 28–43, 2011.

[Solmaz *et al.*, 2012] Berkan Solmaz, Brian E Moore, and Mubarak Shah. Identifying behaviors in crowd scenes using stability analysis for dynamical systems. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(10):2064–2070, 2012.

[Turner and Killian, 1987] Ralph H. Turner and Lewis M. Killian. *Collective Behavior*. Prentice Hall, 1987.

[Van den Berg *et al.*, 2008] J. Van den Berg, Ming Lin, and D. Manocha. Reciprocal velocity obstacles for real-time multi-agent navigation. In *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, pages 1928 –1935, may 2008.

[Zen and Ricci, 2011] G. Zen and E. Ricci. Earth mover's prototypes: A convex learning approach for discovering activity patterns in dynamic scenes. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on*, pages 3225–3232, June 2011.

[Zhong *et al.*, 2015] Jinghui Zhong, Wentong Cai, Linbo Luo, and Haiyan Yin. Learning behavior patterns from video: A data-driven framework for agent-based crowd modeling. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, AAMAS '15, pages 801–809, Richland, SC, 2015. International Foundation for Autonomous Agents and Multiagent Systems.

[Zhou *et al.*, 2012] Bolei Zhou, Xiaogang Wang, and Xiaoou Tang. Understanding collective crowd behaviors: Learning a mixture model of dynamic pedestrian-agents. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 2871–2878, June 2012.