

A Digital Twin City Model for Age-Friendly Communities: Capturing Environmental Distress from Multimodal Sensory Data

Changbum R. Ahn
Texas A&M University
ryanahn@tamu.edu

Youngjib Ham
Texas A&M University
yham@tamu.edu

Jinwoo Kim
Texas A&M University
jwkim@tamu.edu

Jaeyoon Kim
Texas A&M University
skybluego@tamu.edu

Abstract

As the worldwide population is aging, the demands of aging-in-place are also increasing and require smarter and more connected cities to keep mobility independence of older adults. However, today's aging built environment often poses great environmental demands to older adults' mobility and causes their distresses. To better understand and help mitigating older adults' distress in their daily trips, this paper proposes constructing the digital twin city (DTC) model that integrates multimodal data (i.e., physiological sensing, visual sensing) on environmental demands in urban communities, so that such environmental demands can be considered in mobility planning of older adults. Specifically, this paper examines how data acquired from various modalities (i.e., electrodermal activity, gait patterns, visual sensing) can portray environmental demands associated with older adults' mobility. In addition, it discusses the challenges and opportunities of multimodal data fusion in capturing environmental distresses in urban communities.

1. Introduction

Population aging is a global phenomenon. There will be more elderly people than young children in the global population in 2035, and older adults in the US will be over 20% of the entire US population in 2050 [1]. A majority of these elderly people want to live in their present communities as long as possible [2]. Such aging-in-place goals enable older adults to maintain their independence, autonomy, and connection to social support. It offers significant social and economic advantages (e.g., a low rate of clinical depression and obesity) at the individual and societal levels [3]. These goals demand smarter and more connected cities for their independent mobility and healthy aging.

However, there is a growing concern about the inclusiveness and quality of urban built environment

beyond the home; traditional urban planning/design practices do not adequately consider the special needs of older adults experiencing substantial physical and cognitive limitations [4], and overall maintenance conditions of America's aging urban infrastructure (e.g., streets, transit systems, buildings) continue to worsen [5]. Such an urban built environment often presents great physical and/or cognitive demands to older adults' mobility and causes older adults' physical and emotional distress, thereby inhibiting their mobility. Therefore, assessing urban built environment from older adults' perspectives would be the critical first step in fostering age-friendly communities that protect mobility, well-being, and health of older adults.

One promising approach is to exploit new sources of urban data emerging from technological, social, and business innovations. For example, wearable sensing data crowdsourced from urban residents' wearable devices could provide direct observations on how residents feel and respond to the environmental demand of urban communities [6]. In addition, visual sensing data on urban built environment (e.g., Google Street view, crowdsourced photos) could contain key information on environmental stressors in urban communities. Such multimodal data on environmental distress associated with older adults' mobility can be then integrated into the Digital Twin City (DTC) model, in order to provide spatiotemporal patterns of environmental distress associated with older adults' mobility. In this context, this paper aims to examine how data acquired from diverse modalities can be analyzed to represent environmental demands associated with older adults' mobility and can be integrated into a virtual city model. In addition, this paper will discuss challenges and opportunities in fusing such multimodal data, by using experimental data collected in the Bryan downtown, Texas.

2. Research background

2.1 Mobility of older adults

Older adults' mobility plays an important role in their physical well-being, health behaviors, mental wellness, and satisfaction, and has an impact on public health [7]. Many studies indicate that older adults' mobility is closely related to their physical activity [7], social engagement [8], mental disorders [9], nutrition (e.g., consumption of fruit and vegetables) [10], and access to medical services [11]. Protecting and enhancing older adults' mobility is, therefore, the critical first step to promote the quality of their lives.

The mobility of an older adult is mainly affected by his/her physical capacity and surrounding environmental conditions including distance, ambient conditions (e.g., weather conditions), terrain characteristics (e.g., slope, stairs, and uneven surfaces), physical disorders (e.g., vandalized building and litter in the streets), objects or people density in walking path, and speeding cars [3, 5, 6, 7]. Various attempts have been made to assess environmental demands associated with older adults' mobility. Most of them rely on surveys and interviews from actual users (i.e., older adults) and/or trained inspectors [12]. However, opinion surveys and visual audits are not free from subjectivity concerns (i.e., inter-rater reliability issues) and cannot provide continuous assessment for urban communities where many dynamic individual factors (e.g., time of day, traffic, weather, maintenance conditions) interplay [13]. Therefore, a novel breakthrough by leveraging new sources of urban data in understanding the interaction between older adults and surrounding urban built environment is necessary to enable the continuous assessment and create smarter and more connected cities for age-friendly communities.

2.2 Digital twin city (DTC)

The term 'Digital Twin' can be defined as a dynamic digital representation which mirrors and simulates a physical system to help organize and share data for informing better decision-making in the system [14]. To facilitate decision-making regarding the complex systems of urban built environments, many cities have created and leveraged a DTC model based on real-time data from the diverse Internet of Things (IoT) sensors to bridge the gap between real and virtual world [15]. For example, the 'Virtual Rennes' project by the city of Rennes has created the data-rich DTC model to support urban planning and management in the context of the growing population, energy consumption, and environmental issues [16].

In particular, for achieving safe and healthy urban built environment, 'Virtual Singapore' project has created the DTC model informed by IoT sensor-based dynamic data to analyze noise and pollution level in

real-time. Similarly, the DTC models of Singapore have been used by Singapore's security forces to simulate terrorists' attack at a sports stadium in order to mitigate potential risk [17]. The DTC model could be combined with healthcare systems to monitor, diagnose, and predict the health of older adults by integrating the medical physical and virtual spaces; it can provide diverse application scenarios such as real-time supervision, resource optimization and accurate crisis warning systems [18]. As such, the real-time data obtained from monitoring sensors in cities has brought significant potential to improve data-driven decision-making for achieving safe and healthy urban built environments.

2.3 Urban sensing data and environmental distress

New sources of urban data—including infrastructure-based sensors [19], user-generated data [20], and administrative data [21]—are emerging from technological, social, institutional, and business innovations [22]. These new datasets have stimulated empirical data-driven research towards bottom-up sensing of the city [23]. Evaluating urban environments using new urban data to promote community residents' quality of life is one of the areas. Extensive efforts have already been made using passively collected location data from a user's mobile phone [24], crowdsourced self-reports on environmental issues [25], administrative data (e.g., New York City's 311 complaint data) [21], and video and image data (e.g., Google Street view) [26].

One of the promising data sources to better understand human experience in the urban built environment would be physiological responses (e.g., electrodermal activity (EDA), gait patterns, blood volume pulse) captured from wearable devices of older adults. Such physiological responses could provide information regarding how individuals feel and respond to environmental demands, including fight-or-flight responses to threatening stimuli. Chrisinger and King [25], for example, attempted to capture the stress experiences of pedestrians in urban built environments using EDA signals. This research investigated the utility of physiological signals by examining the relationships between subjective evaluation forms (e.g., walkability survey) and physiological signals [26]. Duchowny et al. [27] examines the usefulness of exploiting gait speed and stride length to identify the influence of an environmental demands for mobility. They found variability in accordance with physiological responses in high-demand environments (e.g., absence of traffic signals and sidewalk defects). Also, our previous study [6] highlighted that collective levels of gait stability

and relative heart rates captured from pedestrians in naturalistic ambulatory settings can be indicative of adverse built environmental features that harm neighborhood walkability.

There have been few recent studies on vision-based methods to measure urban perception from images for analyzing changes in the physical appearances in scenes [28,29]. These studies built on the prior works dedicated to analyzing the aesthetic aspects of visual data using generic image features such as color, texture, and SIFT [30-32]. To identify the correlation between visual features and the perceived safety of a Google Street View image, Naik et al. [29] proposed a scene understanding algorithm to rate a perceived safety score, called Streetscore, using training data collected from an online survey with the contributions of participants.

While these studies highlight opportunities to leverage new urban data to capture environmental distress, it has not been fully examined how such data can identify environmental distress to a specific population (e.g., older adults). More importantly, research into whether and how the fusion of multimodal data on urban communities generates added value in capturing environmental distress to older adults is uncharted territory.

3. Capturing environmental distress from urban sensing data

This section discusses a methodology to capture distress (or perceptual distress) from multimodal data sources. In particular, this study focuses on (1) physiological response data captured from older adults' wearable devices, including EDA and gait patterns, and (2) visual sensing data available from various sources (e.g., crowdsourced, Google Street view).

3.1 Capturing environmental distress from wearable sensing data

Although pedestrian's physiological signals have a great potential to understand the effect of the urban built environment, it cannot be directly used in raw data level due to many technical and technological challenges (e.g., motion artifacts and electrode contact noise) [33-36]. Thus, this section will describe an alternative data processing method to effectively capture environmental distress from physiological signals. Additionally, the data processing way was introduced in the author's previous research in detail [37]. The proposed method includes (1) physiological

signal segmentation; (2) extraction of physiological features; (3) calculation of physiological saliency cue (PSC); and (4) PSC aggregation across individuals.

Segmentation is the first step to extract sub-features from physiological signals. Using fixed-length segmentation might not be efficient because signal changes might occur within the analysis segment [33]. Specifically, different physiological signals have different latencies, and each environmental distress has different magnitude and duration of physiological responses [33]. To address these challenges, a non-fixed-length approach is needed, so physiological signals are divided by segments in a data-driven way. Specifically, a bottom-up segmentation is used. It starts with identifying many possible change points and continuously removes less significant ones [38]. The entire signal is partitioned into smaller sub-signals, and then near segments are successively combined by computing similarities between segments. Fig. 1 indicates an example of the bottom-up segmentation of the EDA signal captured in naturalistic ambulatory settings. It shows the usefulness of bottom-up segmentation which captures the signal's appropriate change points to determine distinct segments. The 9 distinct parts were partitioned in the EDA signal using bottom-up segmentation. Each part is visually different from their near neighbors. For example, the first segment in Fig. 1 has low EDA signal values compared to the higher EDA signal values in the second segment. Thus, the bottom-up segmentation was used for both EDA and gait patterns.

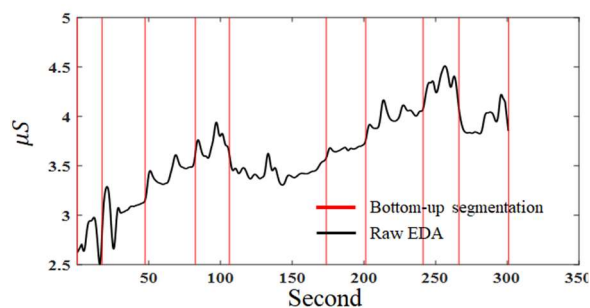


Figure 1. An example of bottom-up segmentation on Electrodermal activity

Secondly, features from each physiological signal (i.e. EDA, gait patterns) were extracted within each segment which is identified by bottom-up segmentation. EDA's features include average skin conductance response (SCR) amplitude, SCR frequency, and mean skin conductance levels (SCL) which are commonly used to portrait emotional arousal [39]. Mean SCL is measured as the mean level

of the EDA signal [39]. Mean SCR amplitude calculates the SCRs' mean amplitude in the time window, and SCR frequency is the number of SCRs partitioned by the length of the time segment [39]. To assess gait patterns, spatiotemporal stride time (ST) is used as a feature. ST is the time span for one heel-strike event and is widely used in ambulatory gait analysis [40]. Specifically, ST is calculated by determining the standard deviation in each segment's gait cycle duration to identify the intensity of the stride-to-stride fluctuations [40].

Thirdly, 'Physiological Saliency Cue' (PSC) was proposed and used to calculate the distinctiveness of one segment in compared to others. The PSC calculation is motivated by an image's contrast calculation in the technique of computer vision [41]. The PSC of segment j for participant i is described by the signal comparison between all other segments of participant i as follows:

$$c_{ij}^{SCL} = \sum_{k=1}^a \frac{t_{ij} u_{ij} - u_{ik}}{T_i \sum_{l=1}^a u_{il}} \quad (1)$$

where u_{ij} is physiological for segment j , T_i is the time duration of the entire signal for i ($i=1, \dots, N$), t_{ij} is the time duration of segment j , and a is the number of segments for participant i . Our previous research highlights the detailed operation process of the PSC equation [37, 42]. The calculated PSC values are normalized in each participant, and then PSC values are computed and aggregated to extract each sub-feature. The aggregation of EDA PSC values is presented by equation 2, and PSC values for gait patterns will be denoted as c_{ij}^{IMU}

$$c_{ij}^{EDA} = c_{ij}^{SCL} + c_{ij}^{SCRFreq} + c_{ij}^{SCRAMP} \quad (2)$$

Lastly, since the physiological response of an individual can be affected by their momentary actions and/or physiological reactivity, collective analysis of physiological signals across many individuals is important to reliably capture environmental distress in each location. For that, histograms of the PSC values are composed for all participants to afford summative information. The diverse histograms indicate different concentration patterns. Specifically, a certain location with high PSC values illustrates a highly right-skewed distribution with negative skewness (See Fig. 2). Fig. 2-b indicates that many individuals presented high PSC values in the location, so its histogram shows skewed towards the right. Therefore, it has a higher negative skewness value (0.51) than Fig. 2a (-1.58).

With these backgrounds, negative skewness is exploited to compute collective PSC values across multiple subjects in one location.

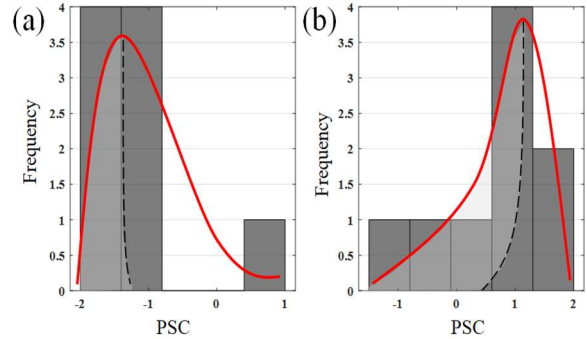


Figure 2. Histogram showing the distribution of physiological saliency cues across all participants; (a) negative skewness value in collective EDA PSC equal to - 1.58; and (b) a negative skewness value in collective EDA PSC equal to + 0.51

3.2 Assessing perceptual distress from visual sensing data

To assess the human perception on scenes that may cause physical and emotional distress, images are ranked through a pairwise comparison. In this paper, participants compare a pair of street-level images to determine which scene looks more stressful or uncomfortable to walk. The outcomes of the pairwise comparison are then converted into a ranked score for each image through the Bayesian graphical model [43]. The ranked score does not have a universal unit as every entity is given a unitless number, and thus that is only effective to compare with others relatively, not deterministically quantify the value [44]. For example, the Microsoft Trueskill used this concept for rating players competing in online games [45, 46] In this paper, building on the Bayesian inference, visual sensing data (i.e., image) is the entity to assess the rank compared with others for understanding the human perception on given scenes in terms of physical and emotional distress, and ultimately predicting that of unexplored paths. In order to rank images through pairwise comparison matches, the visual appearance of streetscapes related with a behavior and health of pedestrians can be collected by crowdsourcing in large city scale [28, 29]. Among diverse matches such as Free-for-All and team games with the different number of players, our experiments were conducted on 1 vs 1 match (i.e., pairwise comparison). Thus, the

algorithm was able to be simplified based on the relationship between two-players.

To cope with the uncertainty in the pairwise comparison, the score of each image is modeled in the form of a Gaussian distribution (i.e., a normal distribution) composed of mean μ (i.e., peak point) and standard deviation σ rather than a single fixed score for each image. In this paper, μ represents the average score of the image; σ represents the degree of uncertainty in the image score. Since the gaussian distribution is characterized that 99.7% of the data are within three times of σ , it was assumed that the initial μ is three times greater than the initial σ in the ranking system. During multiple two-image matches, μ and σ value are iteratively updated: σ value is going to keep lower in conducting every pairwise comparison; the μ of selected image (the winner) is increased, and the μ of not-selected image (the loser) is decreased at every match. Once the pairwise comparison between image A and B is finished, the μ , σ is updated building on [45], for example, in which A wins against B:

$$\mu_A \leftarrow \mu_A + \frac{\sigma_A^2}{c} \cdot v\left(\frac{\mu_A - \mu_B}{c}, \frac{\epsilon}{c}\right) \quad (3)$$

$$\mu_B \leftarrow \mu_B - \frac{\sigma_B^2}{c} \cdot v\left(\frac{\mu_A - \mu_B}{c}, \frac{\epsilon}{c}\right) \quad (4)$$

$$c^2 = 2\beta^2 + \sigma_A^2 + \sigma_B^2 \quad (5)$$

Where μ_A is the mean value of selected image, μ_B is the mean value of not-selected image, σ_A is the standard deviation of selected image, σ_B is the standard deviation of not-selected image, β is the uncertainty due to the performance variation, ϵ is the draw margin. The configurable constant β denotes an uncertainty due to the performance variation, which indicates that β is the additional score above other images to identify an 80% probability of win against the others [43]. ϵ is an empirically estimated value representing the size of draw margin which depends on the probability of draw obtained from empirical tests. The draw margin is a range where the performance of images is assumed to be equivalent even though their values are slightly different. By equation (3) and (4), every image has its own mean value based on the outcome of the pairwise comparison. Accordingly, the standard deviation of each image is changed as follows:

$$\sigma_A^2 \leftarrow \sigma_A^2 \cdot \left(1 - \frac{\sigma_A^2}{c^2} \cdot w\left(\frac{\mu_A - \mu_B}{c}, \frac{\epsilon}{c}\right)\right) \quad (6)$$

$$\sigma_B^2 \leftarrow \sigma_B^2 \cdot \left(1 - \frac{\sigma_B^2}{c^2} \cdot w\left(\frac{\mu_A - \mu_B}{c}, \frac{\epsilon}{c}\right)\right) \quad (7)$$

Here, the function $v(\theta)$ and $w(\theta)$ are defined by the Normal distribution function $N(\theta)$ and Cumulative distribution function $\Phi(\theta)$.

4. Integrating environmental distress into digital twin city: case study

Environmental distress captured from multimodal data sources can be integrated into a DTC model, so that older adults consider such environmental demand in their mobility planning. A case study was conducted to demonstrate the proposed approach to capture environmental distress and integrate into a DTC model.

4.1 Experiment and data collection

Field experiments were performed in the Bryan downtown in Texas to collect data in ambulatory settings. The experiment was performed from June 1st to 2, 2019 between 8 am to 11 am, and the average temperature was 82.99 Fahrenheit (28.33 Celsius). Nine older adults (over 65 years old) were recruited, and EDA and IMU data from the participants were collected while they walked on the pre-defined path with the preferred walking speeds. Wristband-type sensors (Empatica E4) and commercial IMUs (Opal, APDM Inc.) were used for data collection. In addition, Global Positioning System (GPS) data was collected by a smartphone across all participants. EDA and IMU data were sampled at the 4Hz, 125Hz, respectively, and were coordinated with GPS. The Bateaman low-pass filter of 24 sample lengths was used to smooth the EDA signal, and Butterworth low-pass filter with a cut-off 4Hz was used to remove high-frequency noise in the IMU data.

The total distance of the experiment was 1,322.88 ft (403.21 m). Fig. 3 provides visual evidence of terrain rendering of the experiment. In order to recognize the correlation between (1) distress continually captured from physiological responses during the experiment and (2) perceptual distress sparsely measured from visual sensing data (30 images) at a few specific locations, the entire walking path to experiment was divided into 18 POIs (segments) of equal length. Thus, some POIs may contain multiple images; an image which has the lowest score among images was selected. All POIs were used to align participants' PSC values in one POI and image scores. Therefore, these POIs which indicate physical locations are different from the segments by bottom up segmentation.

Several environmental stressors were identified by trained inspectors' assessment and the interviews of participants and included: partially broken wall (POI 5), dumpster (POI 6 and 16), uneven sidewalk (POI 8) (d) dead animal (POI 11), (e) blocked sidewalk by a car (POI 13), and (f) dead branches and leaves overhanging the sidewalk (POI 14).

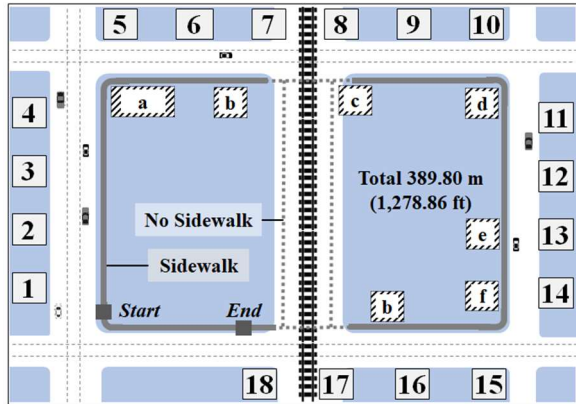


Figure 3. Terrain rendering and environmental distress; (a) partially broken wall, (b) dumpster, (c) uneven sidewalk, (d) dead animals, (e) blocked sidewalk by a car, and (f) dead branches and leaves overhanging the sidewalk

While conducting field experiments to collect EDA and IMU data, a total of 18 geocoded photos were collected along the pre-defined path by using a smartphone (Samsung Galaxy J7). For the pairwise comparison experiments of the collected images, eight participants in 20s to 30s who did not participate in field experiments were instructed to be aware of the question: “which place looks more stressful or uncomfortable to walk?”. To converge the image score for ranking photos, the participants conducted totally 1740 pairwise comparisons including all possible pairs. Each participant did 217 or 218 pairwise comparisons respectively to converge the image score for understanding a level of perceptual distress of each image. The computation time to determine the rank was 0.739s on a computer with the configuration of the Intel i7-8750H CPU and the GTX 1070 8GB GPU. In the proposed approach, the latest score of each image will be only ranked so that the old one will be replaced with the updated one. Here, high rank of an image indicates that the scene looks safe and comfortable to walk.

4.2 Assessment results of environmental distress using physiological signals and visual sensing data

PSC values of EDA and gait patterns were calculated using physiological sensing data. Fig. 4 shows raw physiological signals and PSC values of one subject. As shown in Fig. 4a, the bottom-up segmentation effectively captures the change points of EDA signals, and the resulting PSC values portrait the prominent local patterns. Specifically, the locations which present prominent PSC values include many environmental stressors, such as dumpster (112 to 125 seconds), uneven sidewalk (140 to 150 seconds), and blocked sidewalk by a car (197 to 208 seconds). With respect to gait patterns, PSC values clearly portray distinct patterns around gait cycles 1–118 and 150 to 200, as shown in Fig. 4b. These locations coincide with the POIs containing dead animals, blocked sidewalk by a car, and dead branches and leaves overhanging the sidewalk.

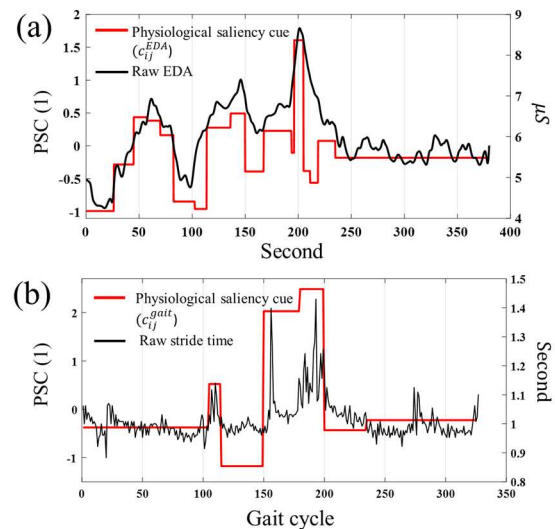


Figure 4. Comparison graph of raw physiological signal and physiological saliency cues; (a) EDA, (b) gait patterns

In addition, perceptual distress of photos taken in the walking path was measured through the pairwise comparisons, as shown in Fig. 5a. In our experiment as shown in Fig. 5b, the pairwise comparisons were conducted by the eight participants described in section 4.1; the μ of each image was eventually converged around 80 comparisons per image approximately.

PSC values of each segment and ranks of visual sensing data were mapped to POIs using GPS coordinates, and PSC values of EDA and gait patterns across all the subjects were aggregated in each POI using the skewness measure. Fig. 6a and 6b indicates that collective PSC values of EDA and gait patterns are correlated with the locations where various

environmental stressors exist. For example, high PSC values of EDA are found in the locations with dumpsters (POI 6), blocked sidewalk by a car (POI 13), and dead branches and leaves overhanging the sidewalk (POI 14), and dead branches and leaves overhanging the sidewalk (POI 14), and collective PSC values of gait patterns (see Fig. 6b) also present high skewness values in several POIs with broken wall (POI 5), uneven sidewalk (POI 8), and dead branches and leaves overhanging the sidewalk (POI 14).

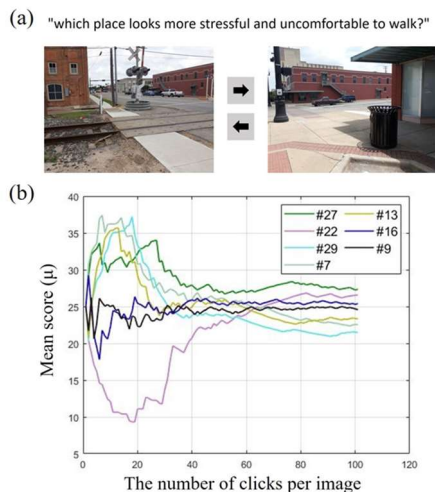


Figure 5. The pairwise image comparisons for ranking score using Bayesian graphical model; (a) an example of pairwise comparison of scenes, (b) results of pairwise comparisons and convergence to a consistent score after 80 clicks in our experiments.

Perceptual distresses captured from visual sensing data are also correlated with the locations having environmental stressors, including uneven sidewalk (POI 8), dead animal (POI 11), blocked sidewalk by a car (POI 13), dead branches and leaves overhanging the sidewalk (POI 14), and dumpster (POI 16).

4.3 Physiological signals and visual sensing data in 3D virtual city model

Visual sensing data and physiological responses obtained from experiments were fed into a 3D virtual city model as shown in Fig. 6e, toward DTC modeling. Objects/areas potentially having a negative impact on older adults' mobility were also imported into the 3D virtual city model through the geospatial localization building upon [46].

Fig. 6e presents objects that may cause physical and emotional distress (e.g., dead animal, uneven

sidewalk, and dumpster) captured by pedestrians in a DTC model. In addition, Fig. 6e illustrates the result of EDA, gait pattern, and the image pairwise comparison at each segment in the DTC model. Using GPS coordinates of the center of each POI, (1) normalized PSC values of EDA, gait patterns and (2) normalized image score were visualized by using different color and size of spheres along with the predefined path. The size of the spheres represents the value of each parameter (i.e., normalized PSC values and image scores) at each POI to help users recognize the correlation between distress captured from physiological responses and perceptual distress measured from visual sensing data. By leveraging smartphone-based physiological signals and crowdsourced visual sensing data reported by pedestrians in the future research, DTC model effectively enables older adults to plan their daily trips based on potential environmental distress in a virtual environment.

5. Opportunities and challenges of multimodal data fusion

The case study results highlight the opportunities to leverage multimodal data in capturing and assessing environmental distress associated with older adults' mobility. Firstly, the results indicate that the effects of several environmental stressors could be manifested in different modality data sources. For example, the segments, including blocked sidewalk by a car (POI 13), and dead branches and leaves overhanging the sidewalk (POI 14), created consistent effects on feature values from all the three sources (i.e., EDA, gait pattern, images). Uneven sidewalk (POI 8) contributed to high distress captured from gait patterns and perceptual distress on visual sensing data. This conforming results from different modalities highlight an opportunity of predicting environmental distress on the locations where a single modality data exist (e.g., only image data is available).

On the other hand, the results indicate that data obtained from multimodalities could be complementary, which would add value in capturing environmental distress. For example, the highest PSC value of EDA was reported at dumpsters (POI 6), while those dumpsters didn't create much effects on the PSC value of gait patterns or perceptual distress on their image. The odors of dumpsters might have impacted the EDA of the participants, while such odors were not visible in their images and did not trigger the change of gait patterns (e.g.,

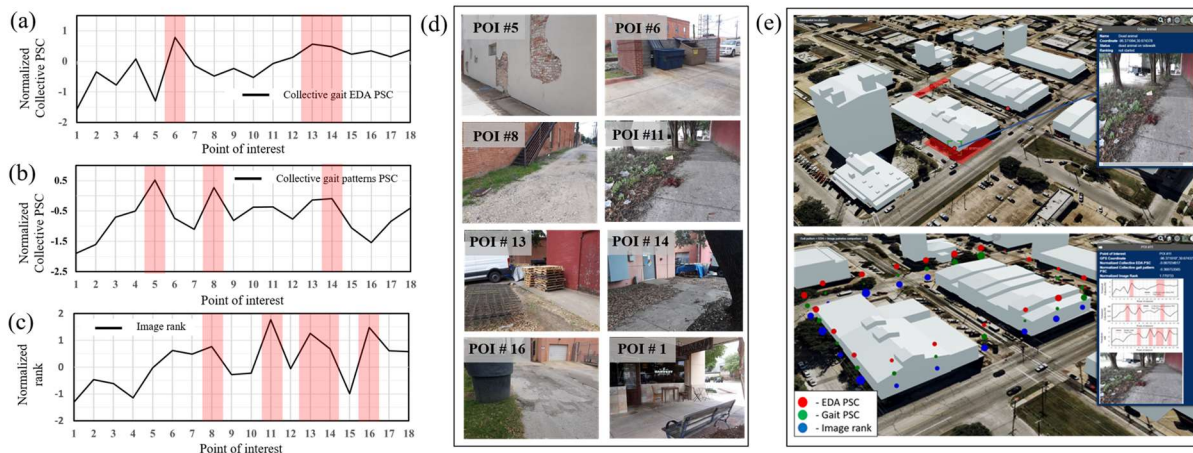


Figure 6. (a) collective PSC for EDA at each POI, (b) collective PSC for gait patterns at each POI, (c) visual sensing data at each POI, (d) photos corresponding high values of POIs, (e) an example of reporting environmental distress by multimodal sensing-based geospatial localization in digital twin city model

avoidance/hesitation actions). These results highlight the potential diversity gained from multimodality. Such diversity potentially enhances the overall performance of capturing stresses and/or the development of the personalized model, compared to the approach using a single modality [47].

However, how to resolve conflicts between different modalities still requires further research. For example, the image of a dead animal body (POI 11) generated the highest perceptual distress, but it did not create much effect to PSC values of EDA and gait patterns; it was found that most the subjects were not able to spot the stressor during their experiments. Likewise, data from different modalities may report on different aspects of the urban built environment, although those data are collected in one location; this may be viewed as no commensurability [44], one of the key challenges in data fusion. Also, there is still a great chance of contradicting results from different modalities, even if data from different modalities deliver observations on the same aspect of the built environment. Future research would be necessary for resolving these issues in leveraging multimodal sensory data.

6. Conclusions

A concept of the DTC model that captures and integrates environmental distress associated with older adults' mobility is proposed by leveraging multimodal data sources data. A case study to demonstrate the concept of the DTC for age-friendly communities was conducted using the experimental data collected in Bryan downtown, Texas. The results highlight that

data captured from various modalities, including physiological sensing data collected from older adults' wearable devices during their daily trips and perceptual distress captured from images on as-is environmental conditions, could identify environmental distress associated with older adults' mobility.

The applications of the proposed DTC model are thus expected to greatly help older adults' mobility planning in a way to identify less-demanding paths considering potential environmental distress in a spatial dimension. However, further efforts are necessary in fusion data obtained from multimodalities, in particular resolving contradicting results from different modalities.

7. Acknowledgements

This study was partially supported by the National Science Foundation–United States (#1800310). Any opinions, findings, conclusions, or recommendations expressed in this article are those of the authors and do not necessarily reflect the views of the National Science Foundation.

8. References

- [1] Ortman, J.M., V.A. Velkoff, and H. Hogan, *An aging nation: the older population in the United States*, United States Census Bureau, Economics and Statistics Administration, US Department of Commerce, 2014.

- [2] Keenan, T.A., *Home and community preferences of the 45+ population*, AARP Research & Strategic Analysis, 2010.
- [3] Wiles, J.L., A. Leibling, N. Guberman, J. Reeve, and R.E.S. Allen, "The Meaning of 'Aging in Place' to Older People", *The Gerontologist* 52(3), 2012, pp. 357–366.
- [4] King, A.C., and J.M. Guralnik, "Maximizing the potential of an aging population", *JAMA* 304(17), 2010, pp. 1944–1945.
- [5] BS, D.A., P.L. PhD, J.P. PhD, T.B. PhD, and I.H.C. MA, "Creating Elder-Friendly Communities", *Journal of Gerontological Social Work* 49(1–2), 2007, pp. 1–18.
- [6] Kim, J., C.R. Ahn, and Y. Nam, "The influence of built environment features on crowdsourced physiological responses of pedestrians in neighborhoods", *Computers, Environment and Urban Systems* 75, 2019, pp. 161–169.
- [7] Clarke, P., and N.A. Gallagher, "Optimizing Mobility in Later Life: The Role of the Urban Built Environment for Older Adults Aging in Place", *Journal of Urban Health* 90(6), 2013, pp. 997–1009.
- [8] Rosso, A.L., J.A. Taylor, L.P. Tabb, and Y.L. Michael, "Mobility, Disability, and Social Engagement in Older Adults", *Journal of Aging and Health* 25(4), 2013, pp. 617–637.
- [9] Lampinen, P., and E. Heikkinen, "Reduced mobility and physical activity as predictors of depressive symptoms among community-dwelling older adults: An eight-year follow-up study", *Aging Clinical and Experimental Research* 15(3), 2003, pp. 205–211.
- [10] Wylie, C., J. Copeman, and S.F.L. Kirk, "Health and social factors affecting the food choice and nutritional intake of elderly people with restricted mobility", *Journal of Human Nutrition and Dietetics* 12(5), 1999, pp. 375–380.
- [11] Glass, T.A., and J.L. Balfour, "Neighborhoods, aging, and functional limitations", *Neighborhoods and health* 1, 2003, pp. 303–334.
- [12] Cunningham, G.O., Y.L. Michael, S.A. Farquhar, and J. Lapidus, "Developing a Reliable Senior Walking Environmental Assessment Tool", *American Journal of Preventive Medicine* 29(3), 2005, pp. 215–217.
- [13] "Urban Planning in 3D: How Creating a Digital Twin Leads to Smarter Cities", *Meeting of the Minds*, 2018. <https://meetingoftheminds.org/urban-planning-3d-creating-digital-twin-leads-smarter-cities-25212>
- [14] "Virtual Singapore", 2019. <https://www.nrf.gov.sg/programmes/virtual-singapore>
- [15] Mohammadi, N., and J.E. Taylor, "Smart city digital twins", *2017 IEEE Symposium Series on Computational Intelligence (SSCI)*, IEEE (2017), 1–5.
- [16] Coletta, C., L. Evans, L. Heaphy, and R. Kitchin, *Creating Smart Cities*, Routledge, 2018.
- [17] Liu, Y., L. Zhang, Y. Yang, et al., "A novel cloud-based framework for the elderly healthcare services using digital twin", *IEEE Access* 7, 2019, pp. 49088–49101.
- [18] Liu, J., Y. Li, M. Chen, W. Dong, and D. Jin, "Software-defined internet of things for smart urban sensing", *IEEE communications magazine* 53(9), 2015, pp. 55–63.
- [19] Zheng, Y., F. Liu, and H.-P. Hsieh, "U-air: When urban air quality inference meets big data", *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, ACM (2013), 1436–1444
- [20] Tasse, D., and J.I. Hong, "Using user-generated content to understand cities", *In Seeing Cities Through Big Data*. Springer, 2017, 49–64.
- [21] Zheng, Y., T. Liu, Y. Wang, Y. Zhu, Y. Liu, and E. Chang, "Diagnosing New York city's noises with ubiquitous data", *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, ACM (2014), 715–725.
- [22] Wang, J., C. Wang, X. Song, and V. Raghavan, "Automatic intersection and traffic rule detection by mining motor-vehicle GPS trajectories", *Computers, Environment and Urban Systems* 64, 2017, pp. 19–29.
- [23] Lathia, N., V. Pejovic, K.K. Rachuri, C. Mascolo, M. Musolesi, and P.J. Rentfrow, "Smartphones for large-scale behavior change interventions", *IEEE Pervasive Computing* 12(3), 2013, pp. 66–73.
- [24] Fan, Z., T. Pei, T. Ma, et al., "Estimation of urban crowd flux based on mobile phone location data: A case study of Beijing, China", *Computers, Environment and Urban Systems* 69, 2018, pp. 114–123
- [25] King, A.C., S.J. Winter, J.L. Sheats, et al., "Leveraging citizen science and information technology for population physical activity promotion", *Translational Journal of the American College of Sports Medicine* 1(4), 2016, pp. 30.
- [26] Wan, L., S. Gao, C. Wu, Y. Jin, M. Mao, and L. Yang, "Big data and urban system model-substitutes or complements? a case study of modelling commuting patterns in beijing", *Computers, Environment and Urban Systems* 68, 2018, pp. 64–77.

- [27] Duchowny, K., P. Clarke, N.A. Gallagher, R. Adams, A.L. Rosso, and N.B. Alexander, "Using Mobile, Wearable, Technology to Understand the Role of Built Environment Demand for Outdoor Mobility", *Environment and Behavior*, 2018, pp. 0013916517749256.
- [28] Dubey, A., N. Naik, D. Parikh, R. Raskar, and C.A. Hidalgo, "Deep learning the city: Quantifying urban perception at a global scale", *European conference on computer vision*, Springer (2016), 196–212.
- [29] Naik, N., J. Philipoom, R. Raskar, and C. Hidalgo, "Streetscore-predicting the perceived safety of one million streetscapes", *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, (2014), 779–785.
- [30] Datta, R., D. Joshi, J. Li, and J.Z. Wang, "Studying aesthetics in photographic images using a computational approach", *European conference on computer vision*, Springer (2006), 288–301.
- [31] Machajdik, J., and A. Hanbury, "Affective image classification using features inspired by psychology and art theory", *Proceedings of the 18th ACM international conference on Multimedia*, ACM (2010), 83–92.
- [32] Marchesotti, L., F. Perronnin, D. Larlus, and G. Csurka, "Assessing the aesthetic quality of photographs using generic image descriptors", *2011 International Conference on Computer Vision*, IEEE (2011), 1784–1791.
- [33] Truong, C., L. Oudre, and N. Vayatis, "Selective review of offline change point detection methods", arXiv:1801.00718 [cs, stat], 2018.
- [34] Choi, B., H. Jebelli, and S. Lee, "Feasibility analysis of electrodermal activity (EDA) acquired from wearable sensors to assess construction workers' perceived risk", *Safety science* 115, 2019, pp. 110–120
- [35] Jebelli, H., B. Choi, H. Kim, and S. Lee, "Feasibility study of a wristband-type wearable sensor to understand construction workers' physical and mental status", *Construction Research Congress*, (2018), 367–377.
- [36] Lee, G., B. Choi, H. Jebelli, C.R. Ahn, and S. Lee, "Reference Signal-Based Method to Remove Respiration Noise in Electrodermal Activity (EDA) Collected from the Field", In *Computing in Civil Engineering 2019: Data, Sensing, and Analytics*. American Society of Civil Engineers Reston, VA, 2019, 17–25.
- [37] Kim, J., M. Yadav, C.R. Ahn, and T. Chaspari, "Saliency Detection Analysis of Pedestrians' Physiological Responses to Assess Adverse Built Environment Features", *The 2019 ASCE International Conference on Computing in Civil Engineering*, ASCE (2019).
- [38] Yang, K., C.R. Ahn, and H. Kim, "Validating ambulatory gait assessment technique for hazard sensing in construction environments", *Automation in Construction* 98, 2019, pp. 302–309.
- [39] Chaspari, T., A. Tsiartas, L.I.S. Duker, S.A. Cermak, and S.S. Narayanan, "EDA-gram: Designing electrodermal activity fingerprints for visualization and feature extraction", *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE (2016), 403–406.
- [40] Hausdorff, J.M., "Gait variability: methods, modeling and meaning", *Journal of NeuroEngineering and Rehabilitation* 2(1), 2005, pp. 19.
- [41] Itti, L., C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis", *IEEE Transactions on Pattern Analysis & Machine Intelligence*(11), 1998, pp. 1254–1259.
- [42] Yadav, M., T. Chaspari, J. Kim, and C.R. Ahn, "Capturing and quantifying emotional distress in the built environment", *Proceedings of the Workshop on Human-Habitat for Health (H3): Human-Habitat Multimodal Interaction for Promoting Health and Well-Being in the Internet of Things Era*, ACM (2018).
- [43] Herbrich, R., T. Minka, and T. Graepel, "TrueSkill™: A Bayesian Skill Rating System", In B. Schölkopf, J.C. Platt and T. Hoffman, eds., *Advances in Neural Information Processing Systems 19*. MIT Press, 2007, 569–576.
- [44] Lahat, D., T. Adali, and C. Jutten, "Multimodal data fusion: an overview of methods, challenges, and prospects", *Proceedings of the IEEE* 103(9), 2015, pp. 1449–1477.
- [45] Moser, J., "Computing Your Skill", 2010. <http://www.moserware.com/2010/03/computing-your-skill.html>
- [46] Kim, J., H. Kim, and Y. Ham, "Mapping Local Vulnerabilities into a 3D City Model through Social Sensing and the CAVE System toward Digital Twin City", *Computing in Civil Engineering* (2019), 451–458.
- [47] Van Mechelen, I., and A.K. Smilde, "A generic linked-mode decomposition model for data fusion", *Chemometrics and Intelligent Laboratory Systems* 104(1), 2010, pp. 83–94.