



Person-Centric Multimedia: How Research Inspirations from Designing Solutions for Individual Users Benefits the Broader Population

Sethuraman Panchanathan, Ramin Tadayon, Hemanth Venkateswara^(✉),
and Troy McDaniel

Center for Cognitive Ubiquitous Computing (CUbic), Arizona State University,
Tempe, USA

{[@asu.edu](mailto:panch,rtadayon,hemanthv,troy.mcdaniel)},
<https://cubic.asu.edu/>

Abstract. While Human-Centered Multimedia Computing (HCMC) improves upon traditional multimedia computing paradigms by accounting for the differences among populations of humans, inter-personal differences between, and intra-personal differences within, populations have created the need for a new paradigm which is sensitive to the needs of a specific user, task and environment. The paradigm of Person-Centered Multimedia Computing (PCMC) addresses this challenge by focusing the design of a system on a single user and challenge, shifting the focus to the individual. It is proposed that this paradigm can then extend the applicability of multimedia technology from the individual user to the broader population through the application of adaptation and integration. These concepts are discussed within the context of disability, where variations among individuals are particularly prevalent. Examples in domain adaptation and autonomous rehabilitative training are presented as proofs-of-concept to illustrate this process within PCMC.

Keywords: Person-centric computing · Coadaptive design
Human-computer interaction · Domain adaptation

1 Introduction

For decades of its existence, multimedia design research has emphasized the development of technology to serve the broader population. In doing so, solutions to challenges at the highest level of user homogeneity had prevailed in favor of encompassing as large a user base as possible. Far from being all-encompassing, however, these solutions have often left many users unable to fully benefit from their use, and still more unable to use them entirely. In perhaps no greater a scenario is this more prevalent than that of users with disabilities; since the focus of design had been on the technology itself rather than the one using it,

those who did not match the ambiguous template of “user”, whether due to the inability to interface with the technology or to do so as effectively as other users who did not share their disability, found themselves at a disadvantage.

To address this fragmentation in the population benefiting from multimedia technology, new computing paradigms have emerged to restore the equilibrium between user and technology in design. The first of these paradigms is that of Human-Centered Multimedia Computing (HCMC) [16]. Human-Centered Multimedia Computing emphasizes the human as the central focus of the process of designing a computing interface. In doing so, this philosophy forces the researcher and designer to be sensitive to the characteristics that define populations of human beings. For example, we differ in our cultures, societal attributes, economies, behaviors, needs, goals and environments. The key advancement of HCMC is the ability for technology to be able to detect, adapt to and accommodate these variations between humans [6]. It is critical to note that this paradigm shift did not occur spontaneously; it is the result of years of advancement in technology and multimedia research, yielding new technologies that have the capability of capturing and utilizing a great deal of information about their users.

Despite the dramatic improvement yielded by HCMC in the applicability of multimedia solutions to a larger variety of individuals, there is still a gap in usability and accessibility that emerges from the fact that in many cases, users of technology vary greatly even at the individual level, and beyond this, even a single individual can change significantly over time. The paradigm of Person-Centered Multimedia Computing (PCMC) addresses this challenge by fully transitioning the focus of design from the technology to the individual [24]. Where HCMC seeks to incorporate inter-societal and inter-population variations in design, PCMC seeks instead to account for inter-personal and even intra-personal variations. The goal of this paradigm is to begin the design process with a challenge posed by a specific individual, yielding a design process wherein the two are integrated at every step. Hence, the technology developed using PCMC guarantees accessibility, usability, and often optimality for an individual. Typically, this fine-tuned design comes at the price of rendering the technology inflexible toward a broader audience. However, it is proposed that individually-inspired design which meets the explicit needs of individuals, using the methods of adaptation and integration, can also meet the implicit needs of a much broader audience. To demonstrate how these two methods expand the impact of the PCMC paradigm beyond what is possible even in HCMC, two example solutions from the Center for Cognitive Ubiquitous Computing are presented, along with the findings in each.

2 Person-Centric Multimedia in Disability Research

The field of disability research is a perfect platform for the application of person-centric design because it is one of the richest sources of intrapersonal variation. Disability is not necessarily a status that divides populations binarily; any individual may be abled in a given context or environment and disabled in another.

For example, in a dark environment, an individual who is blind may actually have a significant advantage in ability over his or her sighted peers due to the heightened sensitivity of the other senses developed as a result of the condition [26]. Hence, disability is largely a matter of the individual, the task and the context, and with this understanding, not only are there variations between individuals but also within a single individual over time. Nevertheless, even by the standard policies used to define disability in today’s society, it affects a significant portion of the world’s population. In the United States alone, an estimated 12.8 percent of the population are classified as individuals with disabilities (www.disabilitystatistics.org). To these individuals, it is of critical importance that an interface is accessible, yet in interfaces designed without the individual nature of disability in mind, accessibility is often nothing more than an afterthought, if it is considered at all [30].

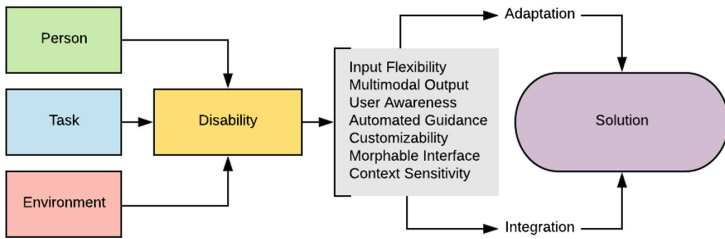


Fig. 1. The process by which person-centric design accounts for disability to provide an accessible, usable interface.

An overview of how person-centric computing weaves accessibility into the design process is given in Fig. 1. Two critical components of the process, adaptation and integration, allow for the applicability of the solution to transition smoothly from context to context. These components are defined and explained below within the example scenarios of domain adaptation and autonomous at-home rehabilitative training.

3 Adaptation: Transition to the Broader Public

Within person-centric computing, it is assumed that with any given technology or interface, it is inevitable that an individual will adapt, or attempt to adapt, to a particular technology over the duration of its use. For example, a new update to an individual’s smart phone interface may change the layout of the menu, requiring that the user learn the latest way to navigate the software on the device. Person centric computing puts into practice the notion that a machine should also be able to detect and respond to variations in a user or context, and to constantly and frequently learn about and adapt to the user in a similar manner to which the user adapts to the machine. As the human and computer

coadapt in this manner, a more reliable and robust interaction between the two is formed [12]. Therefore, when the task, the context, or even the user change, the system designed under PCC principles should evolve in turn to handle this change.

Adaptation in a system can manifest itself in a variety of ways, depending on the intended goal. For example, in the case of accessibility, adaptation might be applied at the input/output level, where the system may, for example, include redundancy of information in feedback across multiple modalities [17] or allow the user to control the way information is exchanged in the interaction [18]. Within a single user, the interface may change in complexity over time as it detects that the user’s skill level or comfort level increases [1]. In addition to users, a system may, for example, adapt to a variety of domains, as presented in the following example of domain adaptation.

4 Example: Domain Adaptation

Data driven models in machine learning do not generalize well across different domains of data. For example, an autonomous driving car trained with data having only sunny-day traffic will fail to perform well when tested during the night. This is on account of the domain difference between the training data and the test environment data. Domain adaptation algorithms transfer knowledge from the source domain to a target domain in the form of models, feature representations and classifiers in order to develop efficient classification models for the target domain. These algorithms play a crucial role in being able to adapt systems across different data distributions. Person-centered algorithms trained using annotated data from a single individual can similarly be adapted to data from a wider population even in the absence of annotations. In this section, we discuss a model for unsupervised domain adaptation that adapts a classifier trained on a source dataset to data from a target domain with a different distribution. We evaluate this adaptation algorithm with face-expression and head-pose estimation datasets.

4.1 Nonlinear Domain Alignment

We first outline the problem of domain adaptation by considering two domains; source domain \mathcal{D}_s and target domain \mathcal{D}_t . The source domain consists of data points $\mathbf{X}_S = [\mathbf{x}_1^s, \dots, \mathbf{x}_{n_s}^s] \in \mathbb{R}^{d \times n_s}$ with associated labels $Y_S = [y_1^s, \dots, y_{n_s}^s]$. The target domain has data points $\mathbf{X}_T = [\mathbf{x}_1^t, \dots, \mathbf{x}_{n_t}^t] \in \mathbb{R}^{d \times n_t}$ and associated labels which are unknown; $Y_T = [y_1^t, \dots, y_{n_t}^t]$. The data points \mathbf{x}_i^s and $\mathbf{x}_i^t \in \mathbb{R}^d$ and the labels y_i^s and $y_i^t \in \{1, \dots, C\}$. We define $\mathbf{X} := [\mathbf{x}_1, \dots, \mathbf{x}_n] = [\mathbf{X}_S, \mathbf{X}_T]$ with $n = n_s + n_t$. The goal is to estimate the labels Y_T under the constraint that the source and target data come from different joint distributions; $P_S(X, Y) \neq P_T(X, Y)$.

A common procedure to align the features of the source and the target is by projecting them to a subspace. The Kernel-PCA (KPCA) estimates a nonlinear basis to account for nonlinear variations in the data. Data is mapped to

an infinite-dimensional subspace defined by $\Phi(\mathbf{X}) = [\phi(\mathbf{x}_1), \dots, \phi(\mathbf{x}_n)]$ before estimating the projection. $\phi : \mathbb{R}^d \rightarrow \mathcal{H}$ is the mapping function and \mathcal{H} is the reproducing kernel hilbert space (RKHS). The similarity between data points \mathbf{x} and \mathbf{y} is represented by the dot product between the mapped representations $k(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x})^\top \phi(\mathbf{y})$. The similarities between all pairs of points is denoted by the kernel matrix $\mathbf{K} = \Phi(\mathbf{X})^\top \Phi(\mathbf{X}) \in \mathbb{R}^{n \times n}$. The kernel matrix is used to estimate the projection matrix \mathbf{A} which is then used to project the data points to a common subspace. The projection matrix \mathbf{A} is estimated by solving,

$$\max_{\mathbf{A}^\top \mathbf{A} = \mathbf{I}} \text{tr}(\mathbf{A}^\top \mathbf{K} \mathbf{H} \mathbf{K}^\top \mathbf{A}), \quad (1)$$

where, \mathbf{H} is the $n \times n$ dimension centering matrix given by $\mathbf{H} = \mathbf{I} - \frac{1}{n} \mathbf{1}$, and \mathbf{I} is an identity matrix and $\mathbf{1}$ is an $n \times n$ dimension matrix of 1s. The projection matrix $\mathbf{A} \in \mathbb{R}^{n \times k}$, is the matrix of coefficients and the nonlinear projected data is given by $\mathbf{Z} = [\mathbf{z}_1, \dots, \mathbf{z}_n] = \mathbf{A}^\top \mathbf{K} \in \mathbb{R}^{k \times n}$.

In order to reduce the source and target domain disparity, we apply Maximum Mean Discrepancy (MMD) [14], to align the projected features of the source and target. Incorporating the MMD, the projection matrix can now be estimated using the formulation,

$$\min_{\mathbf{A}} \left\| \left\| \frac{1}{n_s} \sum_{i=1}^{n_s} \mathbf{A}^\top \mathbf{k}_i - \frac{1}{n_t} \sum_{j=n_s+1}^n \mathbf{A}^\top \mathbf{k}_j \right\|_{\mathcal{H}} \right\|^2 = \text{tr}(\mathbf{A}^\top \mathbf{K} \mathbf{M} \mathbf{K}^\top \mathbf{A}), \quad (2)$$

where \mathbf{M} , is the MMD matrix which is given by,

$$(\mathbf{M})_{ij} = \begin{cases} \frac{1}{n_s n_s}, & \mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}_s \\ \frac{1}{n_t n_t}, & \mathbf{x}_i, \mathbf{x}_j \in \mathcal{D}_t \\ \frac{-1}{n_s n_t}, & \text{otherwise,} \end{cases} \quad (3)$$

Since the goal is to estimate the target data labels, Y_T , we need to estimate a projection where the data points are easily classified. We introduce laplacian eigenmaps to perform similarity based embedding. We cluster data points based on label similarity where data points having the same class label are clustered together ensuring easy classification. Data point similarity is captured using the $(n \times n)$ adjacency matrix \mathbf{W} , where,

$$\mathbf{W}_{ij} := \begin{cases} 1 & y_i^s = y_j^s \text{ or } i = j \\ 0 & y_i^s \neq y_j^s \text{ or labels unknown.} \end{cases} \quad (4)$$

When the sum of squared distances between projected data points (weighted by the adjacency matrix) is minimized, the projected data is clustered by class labels. This is expressed as a minimization problem,

$$\min_{\mathbf{Z}} \frac{1}{2} \sum_{ij} \left\| \frac{\mathbf{z}_i}{\sqrt{d_i}} - \frac{\mathbf{z}_j}{\sqrt{d_j}} \right\|^2 \mathbf{W}_{ij} = \min_{\mathbf{A}} \text{tr}(\mathbf{A}^\top \mathbf{K} \mathbf{L} \mathbf{K}^\top \mathbf{A}). \quad (5)$$

where, \mathbf{L} , denotes the symmetric positive semi-definite graph laplacian matrix with $\mathbf{L} := \mathbf{I} - \mathbf{D}^{-1/2}\mathbf{W}\mathbf{D}^{-1/2}$, \mathbf{I} is an identity matrix and \mathbf{D} , the $(n \times n)$ dimension diagonal matrix with the diagonal entries given by $d_i = \sum_k \mathbf{W}_{ik}$ and $d_j = \sum_k \mathbf{W}_{jk}$. Here, $\|\mathbf{z}_i/\sqrt{d_i} - \mathbf{z}_j/\sqrt{d_j}\|^2$, is the normalized squared Euclidean distance between the projected data points \mathbf{z}_i and \mathbf{z}_j , which are clustered together when $\mathbf{W}_{ij} = 1$, (as they belong to the same category). The projected data is given by $\mathbf{Z} = \mathbf{A}^\top \mathbf{K}$.

4.2 Optimization Problem

We bring together the concepts of projection, domain alignment and similarity embedding to determine the projection matrix in the following optimization problem. Maximizing Eq. (1) while simultaneously minimizing Eqs. (2) and (5) is achieved by maintaining Eq. (1) constant and minimizing Eqs. (2) and (5). The optimization problem is defined as,

$$\min_{\mathbf{A}^\top \mathbf{K} \mathbf{D} \mathbf{K}^\top \mathbf{A} = \mathbf{I}} \text{tr}(\mathbf{A}^\top \mathbf{K} \mathbf{M} \mathbf{K}^\top \mathbf{A}) + \text{tr}(\mathbf{A}^\top \mathbf{K} \mathbf{L} \mathbf{K}^\top \mathbf{A}) + \|\mathbf{A}\|_F^2. \quad (6)$$

where, the last term is the regularizer (Frobenius norm) ensuring a smooth projection matrix. The constraint on \mathbf{A} (in place of equation $\mathbf{A}^\top \mathbf{K} \mathbf{H} \mathbf{K}^\top \mathbf{A} = \mathbf{I}$), is introduced to avoid the projection mapping onto a trivial subspace whose dimensions are less than k , [3]. Equation (6) is solved by introducing the Lagrangian given by,

$$\begin{aligned} L(\mathbf{A}, \Lambda) = & \text{tr}(\mathbf{A}^\top \mathbf{K} \mathbf{M} \mathbf{K}^\top \mathbf{A}) + \text{tr}(\mathbf{A}^\top \mathbf{K} \mathbf{L} \mathbf{K}^\top \mathbf{A}) \\ & + \|\mathbf{A}\|_F^2 + \text{tr}((\mathbf{I} - \mathbf{A}^\top \mathbf{K} \mathbf{D} \mathbf{K}^\top \mathbf{A})\Lambda), \end{aligned} \quad (7)$$

with Lagrangian constants denoted by the diagonal matrix, $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_k)$. Setting the derivative $\frac{\partial L}{\partial \Lambda}$ to 0, we arrive at the generalized eigen-value problem,

$$\left(\mathbf{K} \mathbf{M} \mathbf{K}^\top + \mathbf{K} \mathbf{L} \mathbf{K}^\top + \mathbf{I} \right) \mathbf{A} = \mathbf{K} \mathbf{D} \mathbf{K}^\top \mathbf{A} \Lambda. \quad (8)$$

The k -smallest eigen-vectors of Eq. (8) yield the projection matrix \mathbf{A} . The domain aligned projected features are then obtained by $\mathbf{Z} = \mathbf{A}^\top \mathbf{K}$. We note that a classifier can be trained with only the source data points in the projected data because we only have labels for the source data. However, since the source and target data points are aligned using MMD, we can assume that the source classifier can be used to determine the labels for the projected target data.

4.3 Experiments

In our original work [34], we test our model across different applications like digit recognition, object recognition, facial expression recognition and head-pose recognition. We present only a subset of the results that are relevant for social interaction. At CUbiC, we have spent over a decade designing social assistive aids

for individuals who are blind in order to make social situations more accessible for these users. Our work focuses on vision-based analysis of facial expressions and other visual non-verbal social cues such as gestures and gaze. These cues are extracted from video, captured using a discreet wearable camera, and conveyed to the user through haptics and/or audio. The optimization problem in [34] is also a more generalized version of Eq. (6) and the results are based on the generalized model. For our experiments, we evaluated our model using facial expression recognition and head-pose applications that are relevant for social interaction.

MMI-CKPlus Datasets: The MMI [25], and CKPlus [21] are Facial Expression recognition datasets from which we choose 6 categories of facial expression, viz., *anger*, *disgust*, *fear*, *happy*, *sad* and *surprise*. We generate two domains, CKPlus and MMI, by selecting video frames with the most intense expressions from the two datasets. A pre-trained deep neural network (VGG-F [4]), was used to extract features (4096-dimensional) from the *fc7* layer. PCA was applied to reduce the feature dimensions to 500.

PIE Dataset: The ‘‘Pose, Illumination and Expression’’ (PIE) dataset consists of face images (32×32 pixels) of 68 individuals with varying head-pose, illumination and expression. Along the lines of [20], we create 3 domains based on head-pose, viz. P05 (C05, left pose), P07 (C07, upward pose), P09 (C09, downward pose).

Table 1. Classification accuracies (%) for domain adaptation experiments on facial expression and head pose datasets. {CK+ \rightarrow MMI implies CK+ is source domain and MMI is the target domain. The best results in every experiment (col) are highlighted in **bold**.

Expt	CK+ \rightarrow MMI	MMI \rightarrow CK+	P05 \rightarrow P07	P05 \rightarrow P09	P07 \rightarrow P05	P07 \rightarrow P09	P09 \rightarrow P05	P09 \rightarrow P07	Avg
SA	31.12	39.75	26.64	27.39	25.42	47.24	23.26	41.87	32.84
CA	31.89	37.74	40.33	41.97	41.51	53.43	35.47	47.08	41.18
GFK	28.75	37.94	26.21	27.27	25.27	47.37	21.88	43.09	32.22
TCA	32.72	31.33	40.76	41.79	41.78	51.47	34.69	47.70	40.28
TJM	30.35	40.62	10.80	7.29	16.63	21.69	14.98	27.26	21.20
JDA	29.78	28.39	58.81	54.23	57.62	62.93	50.96	57.95	50.08
NET	29.97	45.83	77.84	70.96	74.55	77.08	73.98	79.01	66.15

In our experiments, we compare the target data recognition accuracies with popular domain adaptation methods such as, Subspace Alignment (SA) [10] Correlation Alignment (CA) [31], Geodesic Flow Kernel (GFK) [13], Transfer Component Analysis (TCA) [23], Transfer Joint Matching (TJM) [19] and Joint Distribution Adaptation (JDA) [20]. Our model is denoted as Nonlinear Embedding Transform (NET). The results we present are based on the generalized model discussed in [34]. The recognition accuracies are depicted in Table 1. The table contains entries for the target data recognition accuracies. The source and

target data are projected to a common domain based on solving Eq. (8). A classifier is trained using the projected source data and the labels for the target data are estimated with this classifier. The NET algorithm provides the best target data classification accuracies compared to the other methods.

The NET algorithm has a straightforward procedure to transfer facial expression recognition and head-pose estimation algorithms across different users. We can apply the NET algorithm to adapt a classification model trained for one user to other users (broader population). This is especially relevant when there is a lack of annotated data in the target dataset. Domain adaptation algorithms help to transfer knowledge from models trained on a population subset to broader populations in the absence of annotated data.

5 Integration: Stealth in Coadaptive Design

While it is of critical importance to design a system which can adapt to a user, it is just as critical to ensure that this adaptation is performed in a way which seamlessly integrates with the task and context under which the user interacts with this system. This integration should account for the complexity of the task and the interests of the user, and attempt to unify these components in a way that minimizes the burden of adaptation of the user to the technology. For example, the user’s task of adapting to haptic feedback from a system designed to convey facial expressions is far less daunting when the haptic patterns are representative of natural facial expressions and directly map to those expressions [2]. When this integration is correctly performed, the adaptation mechanisms of a system are abstracted from the user in such a way that he or she may be completely unaware that they are being used, leading to a state called “stealth adaptation”, named after the concept of “stealth assessment” [28].

In general, the most effective way to integrate stealth adaptation into a system is to find the most natural abstraction of the intended task of an individual given the interaction medium and its constraints and limitations, and to leverage these in design. How complex is the user’s input? What metaphors can be drawn to ease the task of learning on the user? These challenges are presented here in the context of at-home upper extremity motor rehabilitation in the development of a system known as the Autonomous Training Assistant.

5.1 Example: The Autonomous Training Assistant

Overview. The Autonomous Training Assistant (ATA) is a system designed for the automation of guided at-home motor exercise during rehabilitation. The need for this system arose from the explicit need of a single individual, who was hemiparetic as a result of cerebral palsy resulting in one impaired arm and one fully-functioning arm, to complete a series of stick training exercises at home assigned by his martial arts trainer who uses self-defense training as a context for rehabilitative therapy. For this individual, conventional at-home repetitive practice did not provide the in-depth feedback on performance that was provided

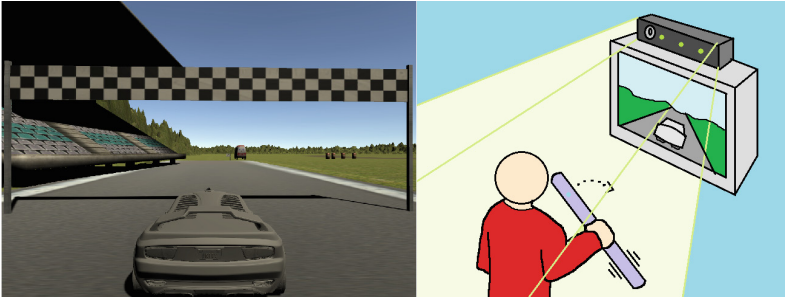


Fig. 2. Overview illustration of the Autonomous Training Assistant system (right) and example game (left).

by the trainer in live exercise sessions between the two individuals, resulting in a less effective learning environment at-home.

Consequently, the two individuals expressed the need for a system that could utilize serious games [9] to deliver a guided, automated exercise experience at home. While this was an explicit need exhibited by this individual, it touches upon a greater challenge in the world of rehabilitation: while frequent exercise in the home is necessary for steady recovery [29], lack of guidance in this environment can lead to reduced compliance to at-home exercise in the long term and, by extension, can slow the progress of the rehabilitation program [27] by dropping the individual outside of the zone of proximal development [35].

A system was developed consisting of a game interface using the Unity platform, a Microsoft Kinect V2 sensor to allow the system to track an individual's joint movements and facial expressions in real-time, and custom exercise equipment, entitled "The Intelligent Stick", equipped with an accelerometer and gyroscope for real-time motion trajectory tracking and haptic motors for vibrotactile feedback on performance as a user attempts an exercise task.

An overview of the ATA system, as originally shown in [32], is illustrated in Fig. 2. To use the system, the user simply picks up the Intelligent Stick and activates the game on Unity. The system then integrates the user's assigned at-home exercises into a game environment where the exercises are used as input to complete the game, and performance in the game is directly reflective of performance at the motion task. Three categories of feedback, as previously derived in [33], are provided to the user. Feedback on posture indicates how correctly the user's body is aligned during exercise, while feedback on progression indicates how accurately the motion itself is performed and pacing indicates how close to the ideal speed the user is moving.

Person-Centric Design. Person-centric design and adaptation principles were employed extensively throughout the creation of the ATA system. The Intelligent Stick device was designed to be modular, allowing for parts to be swapped in and out as necessary. This allows, for example, the use of customized grips for

users with a variety of grip strengths, or various shapes and sizes to match the task and context for each user’s training. The inspiration for this design was that the main user in this project required a custom grip with a strap so that his impaired hand could be secured on the stick during use. It was evident from this stage of the design process that the input mechanism to the system should be flexible enough to account for high individual variability both between individuals who have various tasks and upper extremity strength levels, and within a single individual’s grip strength and exercise types over time.

Furthermore, the mechanism for feedback was designed with the individual in mind. The categories of posture, progression and pacing were derived by directly observing live sessions between the individual and trainer and recording the type of feedback provided by the trainer to the individual during these sessions. The goal of this process was to construct a framework for the evaluation of upper-extremity motor performance that was detailed enough to accurately capture the training program of this particular individual while being broad and flexible enough to encompass some of the most popular standard assessment methods in the field [11,36]. The process of extracting these metrics is detailed in [33].

Finally, the game design process reflected person-centric philosophy by tailoring gameplay specifically to the individual and the task, as described below.

Stealth Adaptation. The approach taken toward stealth integration of adaptation elements in this work draws upon work by Shute et al. entitled “stealth assessment” in academic serious games, where assessment of the subject was performed behind the scenes and integrated directly into the gameplay [28]. To do this, the Evidence Centered Design (ECD) framework is applied [22]. This consists of three models: a task model specifying the type of real task that needs to be performed, a competency model specifying the metrics that indicate competency at that task (in this case, posture, progression and pacing) and an evidence model linking in-game metrics to task performance metrics in the competency model.

In this work, we extend this model to include adaptation. Three popular forms of Dynamic Difficulty Adaptation (DDA) [15] are explored: Bayes net refers to the usage of a Bayesian network to independently track each category of performance and to adapt independently to that category, while hit-rate stabilization refers to the attempt to stabilize the subject at a particular hit-rate, or success rate with respect to the task being performed, and clustering refers to the assignment of a skill level of “high”, “normal” or “low” to a user’s performance based on his or her history of performance at the same motor task. In all three cases, high performance results in an increase in difficulty to maintain challenge while low performance results in a reduction of difficulty.

Evaluation. To evaluate the system for the subject, the ECD framework was employed as the main method of integration. The subject was asked by the trainer to complete a swinging arc motion from the lower right to the upper left of the body. Due to the complexity of this motion, the fruit slicing game pictured

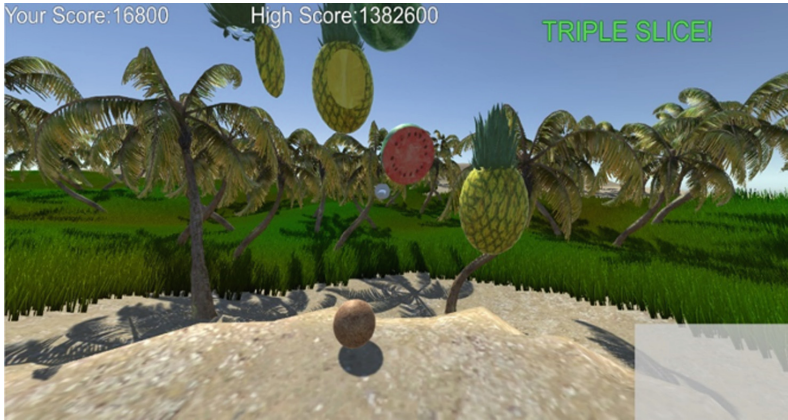


Fig. 3. Screenshot of the Island Fruit game in the case study of the ATA.

in Fig. 3 was developed wherein the sword would be abstracted virtually as a sword intended to slice fruit. Three fruit objects were deployed in the air in each round of a session, and these fruit adapted to the subject's pacing requirement by changing their falling speed, the progression requirement by changing their size, and the postural requirement by being unslicable when the user's impaired arm was off of the intelligent stick. The task of completing the arc motion was thus abstracted to the task of slicing three fruit objects as they fall through the air in virtual space. Three forms of adaptation (Bayes-net, hit-rate stabilization with target hit-rate = 2 fruit sliced, and clustering) and one control condition with no adaptation were experienced by the subject. The subject began with a 1-min tutorial, then played in four 5-min sessions with 10-min breaks inbetween sessions to account for possible learning or fatigue effects.

To determine how well each of the adaptation styles described above fit the needs of the user in this case study, an evaluation was performed using "flow-state", or the ratio in which the subject is considered fully engaged in gameplay to total game time. "Flow" in this context is simply the state in which an individual is given the appropriate level of difficulty in gameplay to match his or her skill level, facing challenges that are neither too difficult (resulting in frustration) or too easy (resulting in boredom) for the player [5]. Flow-state was extracted in this study by leveraging the real-time facial tracking of the Kinect camera. Facial data was fed into a process which would determine the individual's facial expression as a vector of 7 values representing its belief that the user is expressing each of the seven basic emotions contained within the Facial Action Coding System (FACS) [8]. This data was then mapped into a single flow-state output using a method similar to [7]. For each 5-min session, one flow-state value (either "boredom", "anxiety", "flow" or "unknown") was extracted every ten seconds of gameplay for a total of 30 samples in each session.

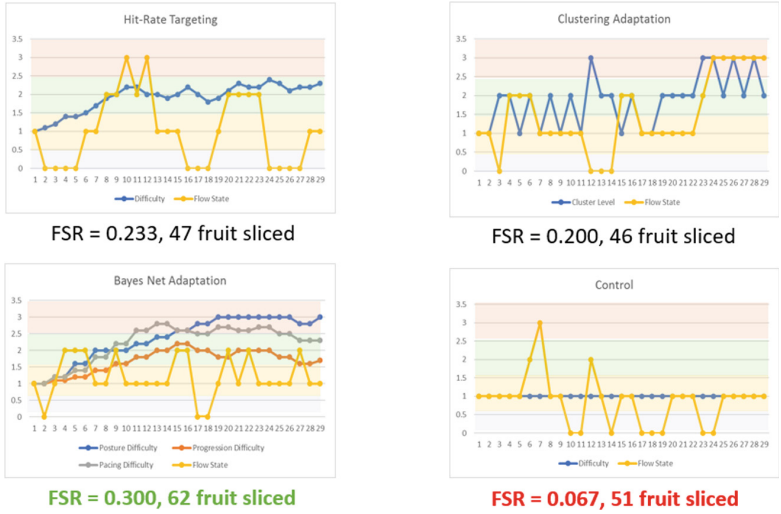


Fig. 4. Flow-state ratio for three adaptation methods (Bayes-net, Hit-rate stabilization, and Clustering) and control condition for Island Fruit game.

Results. Results of the study are shown in Fig. 4. The highest flow-state ratio of 0.300 was yielded in the condition wherein the user was evaluated independently on posture, progression, and pacing using a Bayes net, and each component of performance was then adapted to separately. This finding was also validated by the fact that the subject also demonstrated the highest performance in this condition, and provided the most positive subjective feedback and self-reporting of engagement during gameplay in this particular session. While these results are not reflective of the general effectiveness of these adaptation approaches over larger populations, they have indicated that there is a potential relationship worthy of exploring between the implementation of stealth adaptation in rehabilitative games to the level of engagement and long-term compliance, by extension, of an individual using such a system.

6 Conclusions

The works presented here serve as proofs-of-concept for the highly successful application of person-centric computing within the disability space. With careful design, it is demonstrated that solutions inspired by unique individuals in multimedia need not restrict themselves to those individuals, and can instead benefit the broader population in the same manner that many interfaces which were originally designed to target particular individuals with disabilities have done historically. Under this novel paradigm, individual empowerment fostered by a smart, learning, coadapting interface can redefine the way we view disability in human-computer interaction. Specifically, it can be interpreted in this domain

as a spectrum of ability which can be modulated both by the individual and by the technology to create a space in which any individual is “able” to benefit from interacting with this technology.

Acknowledgements. The authors thank Arizona State University and National Science Foundation for their funding support. This material is partially based upon work supported by the National Science Foundation under Grant Nos. 1069125 and 1116360.

References

1. Afergan, D., et al.: Dynamic difficulty using brain metrics of workload. In: Proceedings of the 32nd Annual ACM Conference on Human Factors in Computing Systems, pp. 3797–3806. ACM (2014)
2. Bala, S., McDaniel, T., Panchanathan, S.: Visual-to-tactile mapping of facial movements for enriched social interactions. In: 2014 IEEE International Symposium on Haptic, Audio and Visual Environments and Games (HAVE), pp. 82–87. IEEE (2014)
3. Belkin, M., Niyogi, P.: Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comput.* **15**(6), 1373–1396 (2003)
4. Chatfield, K., Simonyan, K., Vedaldi, A., Zisserman, A.: Return of the devil in the details: delving deep into convolutional nets. In: BMVC (2014)
5. Chen, J.: Flow in games (and everything else). *Commun. ACM* **50**(4), 31–34 (2007)
6. Cooley, M.: Human-centered design. In: Information Design, pp. 59–81 (2000)
7. Craig, S.D., D’Mello, S., Witherspoon, A., Graesser, A.: Emote aloud during learning with autotutor: applying the facial action coding system to cognitive-affective states during learning. *Cogn. Emot.* **22**(5), 777–788 (2008)
8. Ekman, P., Rosenberg, E.L.: What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS). Oxford University Press, Oxford (1997)
9. Fernandez-Cervantes, V., Stroulia, E., Oliva, L.E., Gonzalez, F., Castillo, C.: Serious games: rehabilitation fuzzy grammar for exercise and therapy compliance. In: 2015 IEEE Games Entertainment Media Conference (GEM), pp. 1–8. IEEE (2015)
10. Fernando, B., Habrard, A., Sebban, M., Tuytelaars, T.: Unsupervised visual domain adaptation using subspace alignment. In: CVPR, pp. 2960–2967 (2013)
11. Fugl-Meyer, A.R., Jääskö, L., Leyman, I., Olsson, S., Steglind, S.: The post-stroke hemiplegic patient. 1. A method for evaluation of physical performance. *Scand. J. Rehabil. Med.* **7**(1), 13–31 (1975)
12. Gallina, P., Bellotto, N., Di Luca, M.: Progressive co-adaptation in human-machine interaction. In: 2015 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO), vol. 2, pp. 362–368. IEEE (2015)
13. Gong, B., Shi, Y., Sha, F., Grauman, K.: Geodesic flow kernel for unsupervised domain adaptation. In: IEEE CVPR (2012)
14. Gretton, A., Smola, A., Huang, J., Schmittfull, M., Borgwardt, K., Schölkopf, B.: Covariate shift by kernel mean matching. In: Dataset Shift in Machine Learning, vol. 3, no. 4, p. 5 (2009)
15. Hunnicke, R.: The case for dynamic difficulty adjustment in games. In: Proceedings of the 2005 ACM SIGCHI International Conference on Advances in Computer Entertainment Technology, pp. 429–433. ACM (2005)

16. Jaimes, A., Sebe, N., Gatica-Perez, D.: Human-centered computing: a multimedia perspective. In: Proceedings of the 14th ACM International Conference on Multimedia, pp. 855–864. ACM (2006)
17. Jewitt, C., Bezemer, J., O'Halloran, K.: *Introducing Multimodality*. Routledge, Abingdon (2016)
18. Jorritsma, W., Cnossen, F., van Ooijen, P.M.: Adaptive support for user interface customization: a study in radiology. *Int. J. Hum.-Comput. Stud.* **77**, 1–9 (2015)
19. Long, M., Wang, J., Ding, G., Sun, J., Yu, P.: Transfer joint matching for unsupervised domain adaptation. In: CVPR, pp. 1410–1417 (2014)
20. Long, M., Wang, J., Ding, G., Sun, J., Yu, P.S.: Transfer feature learning with joint distribution adaptation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2200–2207 (2013)
21. Lucey, P., Cohn, J.F., Kanade, T., Saragih, J., Ambadar, Z., Matthews, I.: The extended Cohn-Kanade dataset (CK+): a complete dataset for action unit and emotion-specified expression. In: CVPR, pp. 94–101. IEEE (2010)
22. Mislevy, R.J., Haertel, G., Riconscente, M., Rutstein, D.W., Ziker, C.: Evidence-centered assessment design. In: Mislevy, R.J., Haertel, G., Riconscente, M., Rutstein, D.W. (eds.) *Assessing Model-Based Reasoning using Evidence-Centered Design*. SS, pp. 19–24. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-52246-3_3
23. Pan, S.J., Tsang, I.W., Kwok, J.T., Yang, Q.: Domain adaptation via transfer component analysis. *IEEE Trans. Neural Netw.* **22**(2), 199–210 (2011)
24. Panchanathan, S., Chakraborty, S., McDaniel, T., Tadayon, R.: Person-centered multimedia computing: a new paradigm inspired by assistive and rehabilitative applications. *IEEE MultiMedia* **23**(3), 12–19 (2016)
25. Pantic, M., Valstar, M., Rademaker, R., Maat, L.: Web-based database for facial expression analysis. In: ICME. IEEE (2005)
26. Rauschecker, J.P.: Compensatory plasticity and sensory substitution in the cerebral cortex. *Trends Neurosci.* **18**(1), 36–43 (1995)
27. Shaughnessy, M., Resnick, B.M., Macko, R.F.: Testing a model of post-stroke exercise behavior. *Rehabil. Nurs.* **31**(1), 15–21 (2006)
28. Shute, V.J., Kim, Y.J.: Formative and stealth assessment. In: Spector, J.M., Merrill, M.D., Elen, J., Bishop, M.J. (eds.) *Handbook of Research on Educational Communications and Technology*, pp. 311–321. Springer, New York (2014). https://doi.org/10.1007/978-1-4614-3185-5_25
29. Smith, D., et al.: Remedial therapy after stroke: a randomised controlled trial. *Br. Med. J. (Clin. Res. Ed.)* **282**(6263), 517–520 (1981)
30. Stephanidis, C.: User interfaces for all: new perspectives into human-computer interaction. In: *User Interfaces for All-Concepts, Methods, and Tools*, vol. 1, pp. 3–17 (2001)
31. Sun, B., Feng, J., Saenko, K.: Return of frustratingly easy domain adaptation. In: ICCV, TASK-CV (2015)
32. Tadayon, R.: *A person-centric design framework for at-home motor learning in serious games*. Ph.D. thesis, Arizona State University (2017)
33. Tadayon, R., et al.: Interactive motor learning with the autonomous training assistant: a case study. In: Kurosu, M. (ed.) *HCI 2015. LNCS*, vol. 9170, pp. 495–506. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-20916-6_46
34. Venkateswara, H., Chakraborty, S., McDaniel, T., Panchanathan, S.: Model selection with nonlinear embedding for unsupervised domain adaptation. In: KnowPros Workshop - Proceedings of the AAAI Conference on Artificial Intelligence (2017)

35. Vygotsky, L.: Zone of proximal development. In: *Mind in Society: The Development of Higher Psychological Processes*, vol. 5291, p. 157 (1987)
36. Wolf, S.L., Catlin, P.A., Ellis, M., Archer, A.L., Morgan, B., Piacentino, A.: Assessing wolf motor function test as outcome measure for research in patients after stroke. *Stroke* **32**(7), 1635–1639 (2001)