Development of An Adaptive Staircase System Actuated by Facial-, Object-, and Voice-Recognition

Alexander Liu Cheng[1,2], Patricio Cruz[3], Wilson Guachamín[4], Carlos Cevallos[4], Benito Ribadeneira[4], Esteban Ortiz[4], Nestor Llorca Vega[2,5]

[1]Faculty of Architecture and the Built Environment, Delft University of Technology, Delft, The Netherlands
[2]Facultad de Arquitectura e Ingenierías, Universidad Internacional SEK, Quito, Ecuador
[3]Facultad de Ingeniería Eléctrica y Electrónica, Escuela Politécnica Nacional, Quito, Ecuador
[4]Facultad de Ingeniería Mecánica, Escuela Politécnica Nacional, Quito, Ecuador
[5]Escuela de Arquitectura, Universidad de Alcalá, Madrid, Spain
E-mail: a.liucheng@tudelft.nl, {patricio.cruz, wilson.guachamin, carlos.cevallos, benito.ribadeneira, esteban.ortiz}@epn.edu.ec, nestor.llorca.arq@uisek.edu.ec

***Abstract*—This paper details a *proof-of-concept* development of an adaptive staircase system-type capable of user-specific mechanical reconfigurations actuated by facial-, object-, and voice-recognition. The system is described via two variation-prototypes—developed at *Technology Readiness Level* 4—as instances of the same system-type. Accordingly, each prototype is informed by the same use-case considerations and requirements. Nevertheless, by means of their mechanical particulars, advantages and disadvantages specific to each variation are identified and explored. The present adaptive staircase system-type consists of two main components, one computational and the other mechanical. The computational component is built upon an inherited System Architecture previously developed and implemented by the authors. More specifically, the computational component uses Google's *TensorFlow* for facial-recognition; *BerryNet* for multi-object detection; and *VoiceIt* for voice-recognition. These three cloud-compatible, -based, or -dependent recognition mechanisms are used to ascertain the identity three user-types: (1) a person without perceivable physical disabilities; (2) a person reliant on a walking-cane; and (3) a person on a wheelchair. With the exception of the first case, the computational component proceeds to actuate mechanical transformations pertinent to each variety of disabilities depending on which user-type is identified. The objective of this implementations is to present an intuitive and automated vertical mobility solution capable of supporting users with varying degrees of reduced mobility.**

***Keywords*—Internet of Things, Ambient Intelligence, Active and Assisted Living, Wireless Sensor and Actuator Networks, Adaptive Staircase, Adaptive Architecture**

I. Introduction and Background

This paper details a *proof-of-concept* development of an adaptive staircase system-type, which is characterized by two core features: its capabilities (1) to adapt mechanically to accommodate to and/or to compensate for the reduced mobility associated with physical disabilities; and (2) to engage said adaptation under specific recognized conditions as enabled by facial-, object-, and voice-recognition via cloud-based or -dependent *Machine Learning* (ML).

Two variations of this system-type are developed, with both able to recognize any of the following three user-types: (i) a person without perceivable physical disabilities; (ii) a person reliant on a walking-cane; and (iii) a person on a wheelchair. The first user-type does not instigate actuation; it is deliberately considered to account for the system's ability to recognize the absence of explicit physical disabilities. The second user-type instigates physical changes in the dimensions of the stair's tread and/or riser to facilitate more effortless stair-climbing. The third user-type instigates the transformation of the staircase into an elevating platform, which dismisses stair-climbing entirely. Both staircase variations have the same reaction to the first and third user-types. With respect to reactions to the second user-type, the first staircase variation (see Section II.A and Section III) adjusts the dimensions of *both* its tread and riser (see Figure 1) while the second variation (see Section II.B and Section III) can only adjust its riser (see Figure 2). Nevertheless, due to its mechanical design, the second variation enables the second user-type to choose to engage in partial stair-climbing or to avoid it entirely; whereas the first variation requires stair-climbing. The reasoning behind each variation is explained in *Concept and Approach* in Section II, and their advantages as well as disadvantages are discussed in *Results and Conclusions* in Section IV. Both variations aim to empower users with varying physical disabilities to ascertain a degree of independence with respect to mobility, which promotes dignity and quality of life.

The present work situates the development of an adaptive staircase system—via the implementation of two variation-instances—within a comprehensive *intelligent built-environment* discourse. That is, a discourse that is informed by *both* (I) *Ambient Intelligence* (AmI) [1], *Ambient Assisted Living*—or *Active and Assisted Living*—(AAL) [2] with respect to *Information and Communication Technologies* (ICTs); and (II) *Adaptive Architecture / Interactive Architecture* [3] with respect to *Architecture, Engineering, and Construction* (AEC) considerations.

II. Concept and Approach

The prototypes of both staircase-variations are built at 1:2.5 scale and at *Technology Readiness Level* (TRL) 4 [4]. Each prototype is driven by a *microcontroller unit* (MCU) attached to a variety of sensors, actuators, and emitters. In the present implementations, all sensors and emitters respond to user-safety considerations by either preventing actuation or by providing notifications via sound as well as light emission. Such safety measures ensure that actuation does not take place while the user is climbing steps (while in stairs-mode), nor while the user is outside of the bounds of the elevating platform (while in elevator-mode). Each variation's MCU communicates with a coordinating single-board computer used to operate the facial-, object-, and voice-recognition mechanisms. In the present setup, the single-board computers are already trained to recognize a variety of faces (see Figure 3) and associated voices (see Figure 4) as corresponding to particular user-types. Similarly, the object-recognition mechanism is already pre-trained to recognize wheelchairs as well as other assistive devices. In both variations, facial-recognition is first engaged to detect faces who approach the system to within a meter. If the identity of the person is ascertained to a confidence level greater than 70%, the system actuates to its corresponding configuration. If, however, the confidence level is low, the user is prompted to utter a predetermined phrase. The voice-recognition mechanism detects both the phrase as well as the identity of the person uttering it. Actuation is engaged by correlating both the facial- and voice-recognition output. Finally, if both the facial- and voice-recognition mechanisms failed to ascertained the identity of a user, the remaining object-recognition mechanism attempts to recognize a wheelchair, walking-cane, etc., and actuates accordingly.

A. Staircase variation 1

The first staircase variation features three mechanical modes / configurations (see Figure 1) corresponding to each of the three user-types mentioned in Section I. The first mode represents a staircase that is compliant with both the *Ecuadorian Service for Standardization* (INEN) prescriptions [5] as well as with the American *Occupational Safety and Health Administration* (OSHA) standards [6]. That is, the staircase's steps are dimensioned with 30 cm treads and 17 cm risers, which may be comfortably climbed by users without physical disabilities. The second mode represents a staircase that is designed for people with mild to moderate physical disabilities with respect to mobility. That is to say, the step's tread width expands to 42 cm while its riser height decreases to 8.5 cm. This second mode is intended for the elderly as well as for pregnant women / nursing parents whose mobility may be reduced. The third and final mode represents an elevating platform that eliminates the need to climb steps entirely. This mode is designed for people dependent on wheelchairs and/or on other mobility support-devices such as rollators, walkers, etc.—i.e., for people whose physical ability to climb steps is either impossible or unduly difficult.

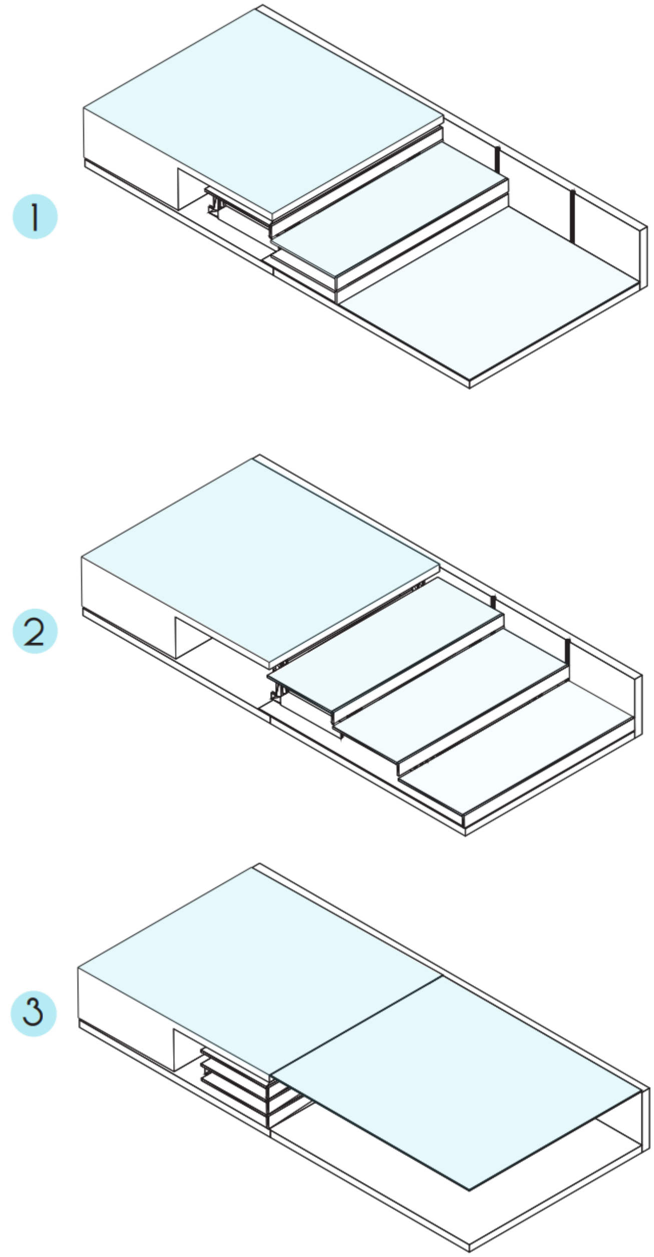

Figure 1. Staircase variation 1 reacting to (1) the first user-type (tread depth: 30 cm; riser height: 17 cm); (2) the second user-type (tread: 42 cm; riser: 8.5 cm); and (3) the third user-type (steps are retracted to enable platform elevation).

B. Staircase variation 2

The second staircase variation also features three mechanical modes / configurations (see Figure 2). The first mode, as in staircase variation 1's first mode, represents an INEN / OSHA compliant staircase. The second mode is capable of instantiating multiple riser heights, which may be used in cases of rehabilitation, where the user is encouraged to walk or to train in a variety of climbing heights. In this mode, a user undergoing rehabilitation may use the system to gradually increase her stair-climbing ability over time.

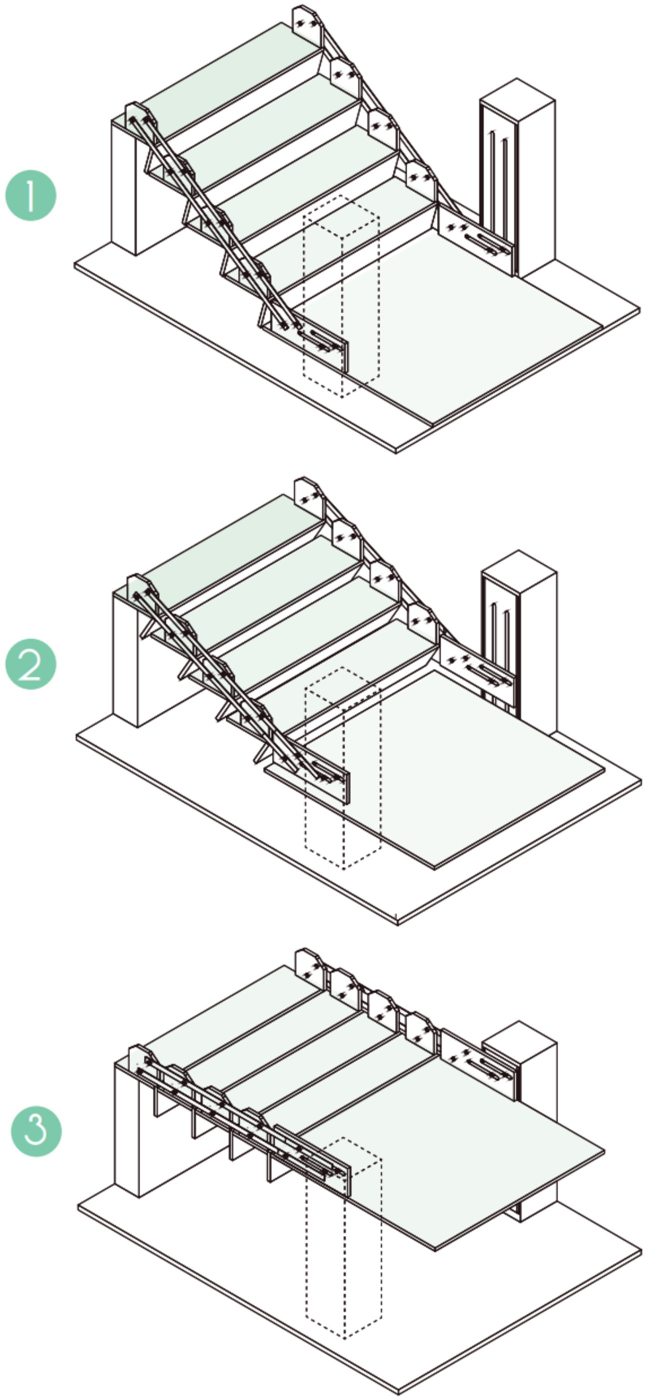

Figure 2. Staircase variation 2 reacting to (1) the first user-type (with standards-compliant tread-vs.-riser dimensions and proportions); (2) the second user-type—or to users with a variety of disabilities in state of rehabilitation; and (3) the third user-type (the landing platform elevates).

The third mode turns the staircase's landing into an elevating platform. In this variation, the first mode caters to the first user-type (user without physical disabilities), while the third mode to the second and third user-types (users dependent on canes, walkers, rollators, or wheelchairs). The second mode principally caters to users in rehabilitation.

III. Methodology and Implementation

The staircase system-type consists of two main components, one computational and the other mechanical. The computational consists of three mechanisms, each concerned with facial-, object-, and voice-recognition. The mechanical consists of the physical parts that instantiate the reconfiguration modes particular to each user-type. Each of the computational mechanisms inherits and/or builds upon previous developments via *Application Programming Interfaces* (APIs) by the authors (with respect to facial- [7], object- [8], and voice-recognition [9]). That is to say:

With respect to the computational mechanisms, the facial-recognition mechanism is developed via Google's free and open-source *TensorFlow* [10]. Its functionality is implemented in Python on the single-board computer, which enables it—with the assistance of a low-cost *Raspberry Pi Camera Module V2*—to recognize faces via cloud-based ML. The face of each of the test-individuals is captured in various positions and used to generate a profile (see Figure 3). In this implementation, each profile is associated with a user-type. The object-recognition mechanism is developed via *BerryNet* [11], which is built with *Inception* ver. 3 [12]) for a classification model and with *TinyYOLO* [13] for a detection model. *BerryNet* serves as a localized *Deep Learning* gateway implementable on a single-board computer, although its performance—as well as that of *TensorFlow*'s implementation—benefit from a cluster of said computers rather than in a single instance.

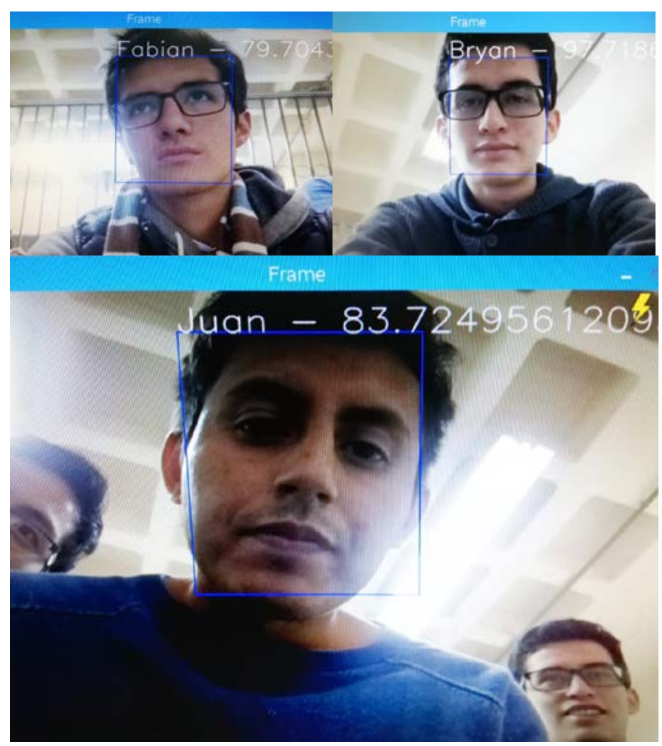

Figure 3. Three sample faces used to associate identity with user-type and to enable facial-recognition and corresponding system-actuation.

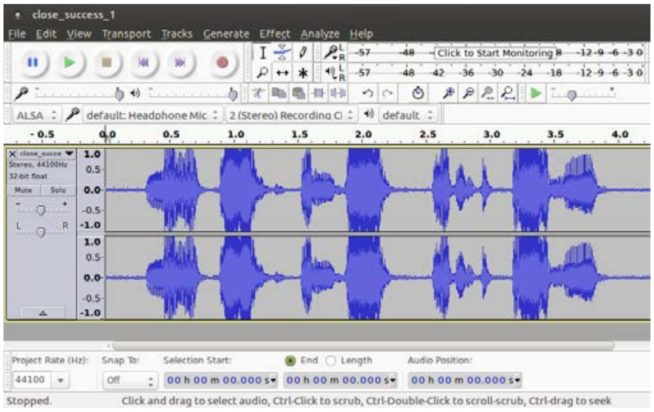

Figure 4. VoiceIt Authentication Success in close-range, as previously implemented by the authors [9].

The voice-recognition mechanism is built via *VoiceIt*'s [14] API in Python. User-profiles are created for each test-subject associated with a user-type. In this process, a minimum of three voice samples are required to *enroll* each subject. Following successful enrollment, the identity of each subject may be ascertained via his/her voice (see Figure 4). This mechanism is capable of recognizing both the utterance as well as the identity of the utterer. In the present implementation only the identity of the utterer is used.

The above three ML-based recognition mechanisms are implemented in order to correlate their outputs to increase the accuracy of identity-detection. In instances where the camera's visibility is unhindered and recognition confidence is greater than 90%, the facial-recognition mechanism takes precedence. In instances where the facial-recognition confidence is greater than 70% yet lesser than 90%, the identity of the user is ascertained via a correlation of the facial-recognition mechanism's and voice-recognition mechanism's output. In cases where facial-recognition confidence is below 70%, the voice-recognition mechanism takes precedence. Finally, in cases where both facial- and voice-recognition failed, the object-recognition mechanism takes precedence—that is, perhaps the user is someone whose profile is not yet stored in the system but is nevertheless recognized to be on a wheelchair or using a rollator, walker, etc.

With respect to the mechanical component, the first staircase variation is built with MDF. Its retracting / extending function is driven by two stepper motors (see Figure 5, 1 and 2), while its platform's elevating function is driven by four stepper motors (with corresponding drivers) (see Figure 5, 3-5). The second staircase variation is built with aluminum parts. Its mechanical transformation is enabled by two stepper motors (with corresponding drivers) built into the support rails (see Figure 6). As both variations are built at 1:2.5, the motors that actuate the system are not rated for real-scale use, but are appropriate for the present *proof-of-concept* implementations. Moreover, in both cases the camera is detached from the actual prototype and is situated adjacent to it for practicality during tests and trials.

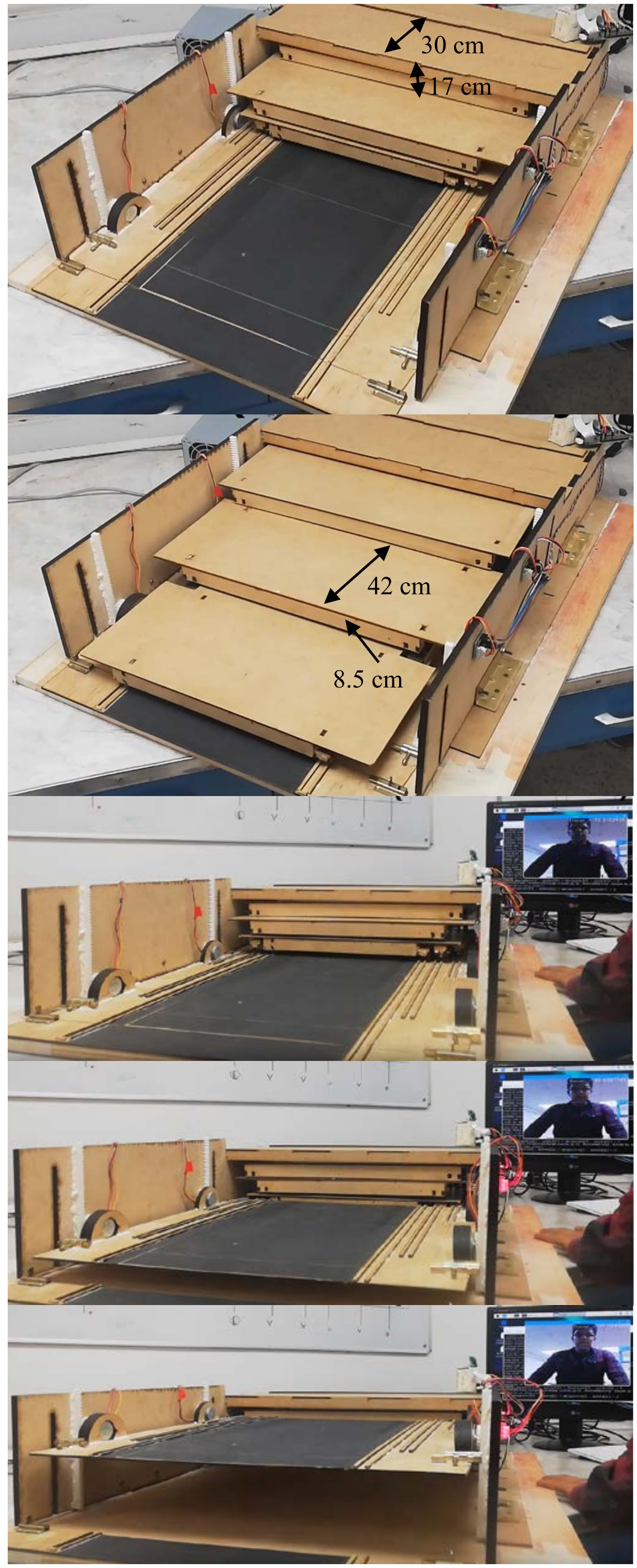

Figure 5. From top-to-bottom: staircase variation 1 configured for (1) the first user-type (standard-stairs); (2) the second user-type (easy-stairs); and (3-5) the third user-type (elevating platform).

Figure 6. Staircase variation 2 configuring for (Top and Bottom) the first user-type (standard-stairs); and (Middle) the third user-type (elevating platform).

IV. Results and Conclusions

The present staircase variations of the same system-type performed as expected in every test. That is to say, all user-types were recognized via predefined test-subject profiles, and mechanical transformations were consistently pertinent to each user-type's presupposed physical disability. Moreover, with respect to staircase variation 1 (see Figure 1 and Figure 5), the speeds of deployment for the configurations corresponding to the three user-types were as follows. For the first user-type (standard building-code compliant stairs, i.e.: 30 cm tread depth by 17 cm riser height): 20 seconds (± 0.5 seconds) to extend; 21 seconds (± 0.5 seconds) to retract. For the second user-type (easy-stairs, i.e.: 42 cm tread depth by 8.5 cm riser height): 30 seconds (± 0.8 seconds) to extend; 32 seconds (± 0.2 seconds) to retract. Finally, for the third user-type (engaging the platform): 4 seconds (± 0.2 seconds) for the platform to elevate; 3.5 seconds (± 0.1 seconds) to collapse—N.B.: previous staircase-configuration deployment times must be added to platform elevation times accordingly, depending on which configuration was instantiated immediately before system recognized a third user-type and began retracting the steps. With respect to staircase variation 2: since the system engaged a single vertical transformation to address all user-types, only the speeds of full platform elevation and collapsion were measured: 70 seconds (± 0.5 seconds) and 64 seconds (± 0.3 seconds), respectively. From these values other extension / retraction degrees may be estimated. In both staircase-configurations, the deviation with respect to expected deployment times were within acceptable margins. Such deviations were principally due to expected degrees of friction, degrees which would need to be reconsidered before the variations may be developed further to a higher TRL.

While the variations operated as planned, their performance was hindered by a number of limitations. First and foremost, the limited computational resources available in each variation's single-board computer hindered real-time facial- and object-recognition. The authors have previously solved this issue by creating clusters of single-board computers in order to instantiate a decentralized super computer. However, while this approach may be useful as a back-up and local solution, the principal approach should make use of cloud-based services with considerably more powerful computational resources. The local approach has its place in the present System Architecture, yet an increasing amount of ML mechanisms featured in Intelligent Built-Environment implementations make the local cluster-model untenable due to the sheer amount of computational resources required to run all said ML mechanisms simultaneously. Instead of doing so, the local cluster should be used to run mechanisms selectively, when the corresponding cloud-based resources are unavailable. Accordingly, and supposing a local cluster approach as back-up, future iterations of the present staircase system-type should implement its facial-recognition mechanism via services such as Google's *Cloud TensorFlow Processing Unit* (TPU) [15]. In the present implementation facial- and object-recognition were undertaken by different mechanisms (TensforFlow vs. BerryNet). This was in consideration of the limited resources of the single-board computer, where these mechanisms were never executed simultaneously. But once Cloud TPU is implemented, both facial- and object-recognition will be undertaken by the same mechanism.

Another limitation in the present implementation concerned the use of *VoiceIt*'s capabilities. At present, only the identity-via-voice feature was implemented, while its speech-to-text capability was not. However, it may be useful for the present staircase variations to be able to accept verbal commands from particular users, and not only to be able to identify such users. For example, a user may verbally command an actuation to stop in unforeseen situations, or he/she may command the systems from distances greater than ~one meter (the present implementation's operation distance-limitation for the camera to engage in facial- or object-recognition).

Finally, at present real-scale iterations of both systems are being built with the above-mentioned considerations as well as with motors and parts appropriate for human interaction at said scale.

ACKNOWLEDGEMENTS

The authors acknowledge the assistance of the following students from *Facultad de Ingeniería Electrica y Electrónica* (FIEE) as well as *Facultad de Ingeniería Mecánica* (FIM) at *Escuela Politécnica Nacional* (EPN): Byron Ayala, Juan Balseca, Daysi Chávez, José Correa, Daniel Córdoba, Esteban Cúñez, Ariel Herrera, Denisse Leines, Javier Mateus, Kevin Morejón, Fabián Poveda, Alejandra Quishpe, Johana Quiñónez, Bryan Sagnay, Gustavo Valencia, Santiago Vázques.

Moreover, the authors acknowledge the support provided by Professors Jackelin Abad and Ivan Chico at FIEE, EPN; Víctor Hidalgo at FIM, EPN; and Nancy Carchipulla at *Facultad de Arquitectura e Ingenierías* (FAI)*, Universidad Internacional SEK* (UISEK). Finally, the authors acknowledge Jorge Jarrín and Francisco Cevallos at *Estudio 685* as well as Diego Bermeo at FAI, UISEK for their assistance in coordination and production.

Part of the work detailed in this paper is made possible by funding provided by EPN's Project *PVS-2018-032* and *PVS-2018-038* as well as UISEK's Project *P111819*.

REFERENCES

[1] A. Kameas and K. Stathis, Eds, *Ambient Intelligence: 14th European Conference, AmI 2018, Larnaca, Cyprus, November 12-14, 2018. Proceedings,* 1st ed. Cham: Springer International Publishing; Springer, 2018.

[2] D. Calvaresi *et al,* "Exploring the ambient assisted living domain: A systematic review," *Journal of Ambient Intelligence and Humanized Computing*, vol. 8, no. 2, pp. 239–257, 2017.

[3] M. Fox, *Interactive architecture: Adaptive world*. New York: Princeton Architectural Press, 2016.

[4] European Association of Research and Technology Organisations (EARTO), *The TRL Scale as a Research & Innovation Policy TOOL: EARTO Recommendations.* [Online] Available: http://www.earto.eu/fileadmin/content/03_Publications/The_TRL_Scale_as_a_R_I_Policy_Tool_-_EARTO_Recommendations_-_Final.pdf. Accessed on: Jan. 07 2015.

[5] *NTE INEN 2249: Accesibilidad de las personas al medio físico. Circulaciones Verticales. Escaleras. Requisitos,* 2249, 2016.

[6] *United States Department of Labor, Occupational Safety and Health Administration, Part 1910: Stairways,* 1910.25.

[7] A. Liu Cheng, H. Bier, and G. Latorre, "Actuation Confirmation and Negation via Facial-Identity and -Expression Recognition," in *Proceedings of the 3rd IEEE Ecuador Technical Chapters Meeting (ETCM) 2018*, 2018.

[8] A. Liu Cheng, H. Bier, and S. Mostafavi, "Deep Learning Object-Recognition in a Design-to-Robotic-Production and -Operation Implementation," in *Proceedings of the 2nd IEEE Ecuador Technical Chapters Meeting 2017*, Guayaquil, Ecuador, 2017.

[9] C. Follini, A. Liu Cheng, G. Latorre, and L. Freire Amores, "Design and Development of a Novel Robotic Gripper for Automated Scaffolding Assembly," in *Proceedings of the 3rd IEEE Ecuador Technical Chapters Meeting (ETCM) 2018*, 2018.

[10] TensorFlow™, *An open source machine learning framework for everyone.* [Online] Available: https://www.tensorflow.org/. Accessed on: Apr. 20 2018.

[11] DT42©, Ltd, *BerryNet®: Deep learning gateway on Raspberry Pi.* [Online] Available: https://github.com/DT42/BerryNet. Accessed on: Jun. 21 2017.

[12] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, *Rethinking the Inception Architecture for Computer Vision.* Available: http://arxiv.org/pdf/1512.00567.

[13] J. Redmon and A. Farhadi, *YOLO9000: Better, Faster, Stronger.* Available: http://arxiv.org/pdf/1612.08242.

[14] VoiceIt®, *Your Voice Is Key®: Cloud Based Biometrics Service.* [Online] Available: https://voiceit.io/. Accessed on: May 30 2018.

[15] Google Cloud Platform®, *Cloud TPU: Train and run machine learning models faster than ever before.* [Online] Available: https://cloud.google.com/tpu/. Accessed on: May 03 2019.