Stroke Modeling and Synthesis for Robotic and Virtual Patient Simulators

Maryam Pourebadi[1] **and Laurel D. Riek**[2]*

Computer Science and Engineering Department
University of California
San Diego, California 92122
[1]pourebadi@ucsd.edu
[2]lriek@ucsd.edu

Abstract

Stroke affects 15 million people each year, and is the fourth most common misdiagnosis reported by clinicians. Research shows using robotic and virtual patient simulators may reduce preventable patient harm. However, current commercial patient simulators have static faces and lack the realistic depiction of non-verbal facial cues important for rapid diagnosis of neurological emergencies such as stroke. Their lack of expressions may adversely affect clinical learners' learning performance and contribute to a failure to recognize a stroke in their future careers. Our multidisciplinary research addresses the urgent need for an expressive training tool by developing techniques to automatically model the facial characteristics of neurological impairment in real-time, and use the model to transfer these onto patient simulators with different genders, races, and ages. By leveraging the use of these expressive patient simulators, clinical learners will have the potential to more accurately diagnose people with stroke and other neurological impairments, increase educational performance, and improve patient interaction when interacting with patients from diverse backgrounds.

Introduction

Stroke, a substantial contributor to the global disease burden, affects 15 million people each year, and is the second leading cause of death worldwide [19, 5]. One of the contributors to this disease burden is diagnostic failure: stroke is the fourth most common misdiagnosis reported by clinicians [20]. This may be a result of the fact that clinical learners (CLs) often fail to achieve mastery of the neurological examination when training on patient simulators [7]. More effective clinical training in the delivery of the neurological exam and interpretation of its findings may help improve clinicians' confidence in their exams, increase the accuracy of impairment diagnosis, and reduce the incidence of preventable patient harm [14, 8].

Simulation-based clinical learning (SBL) is an important component of clinical education that is already incorporated into several subspecialties that intersect with treating stroke [18, 16]. SBL provides CLs with low-risk, high-fidelity, clinically-similar learning environments which can simulate a range of scenarios. SBL enables CLs to safely practice their clinical and procedural skills without the fear of harming real patients [15]. (See Fig. 1.) Three main tools used in clinical education are robotic patient simulators (RPS), augmented reality simulators (ARS), and virtual patient simulators (VPS). RPSs are lifelike android robots that can simulate realistic patient physiologies and pathologies (See Fig. 2, left). ARSs are simulators that use augmented reality techniques to combine physical human-shaped surfaces with dynamic visual imagery projected on its surface (See Fig. 2, center). VPSs are interactive digital simulations of real patients, displayed on a monitor (See Fig. 2, right).

There are many benefits to SBL. Research shows using patient simulators may reduce preventable patient harm, which cause nearly 400,000 deaths per year in US hospitals alone [8]. Another study suggests that using these patient simulators has a positive influence on CLs' comprehension, efficiency, and enthusiasm for learning [9]. Furthermore, researchers found that when compared with non-digital educational methods, patient simulator systems are superior in terms of improving knowledge, and skill-building [11].

Despite the benefits of using RPS, ARS, and VPS systems, current commercial systems have static faces and lack the realistic depiction of non-verbal facial cues important for rapid diagnosis of neurological emergencies such as stroke [17]. This is problematic because facial expressions (FEs) serve as an important social function and clinical cue in patients. Therefore, the lack of expressions in simulators may adversely affect CLs' learning performance and failure to recognize a stroke in their future careers.

While some researchers are working on developing expressive RPS/APS/VPS simulators, one challenge is how to create a generalized model of a neurological impairment based on data from a limited group of patients. Addressing this challenge can lead to development of generalized expressive simulator systems which are capable of representing a diverse group of patients with different ages, genders, and ethnicities. It can ultimately enhance healthcare education with realistic, expressive simulators capable of mimick-

*This material is based upon work supported by the National Science Foundation under Grant Nos. IIS-1253935, CMMI-1935500.
Copyright © 2020, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

Figure 1: Robotic patient simulators are a unique application space in HRI. They are tele-operated, life-size mannequins that can exhibit over 5000 physiological variables, and can breathe, bleed, and respond to medications. However, they are largely inexpressive, leading to poor immersion and poor training outcomes for CLs, and possibly poor clinical outcomes for patients. Our work addresses this gap by introducing patient simulator systems with a much wider range of expressivity, including the ability to express pain, neurological impairment, and other pathologies, across simulators with diverse genders, races, and ages.

ing patient-like FEs. This has the potential to positively affect CLs' retention and revolutionize healthcare education.

Thus, our work examines state-of-the-art technical approaches for human facial expression analysis, facial action modeling, and facial expression synthesis on RPS/ARS/VPS systems. We contextualize this work within the field of human-robot interaction (HRI), as ultimately we are interested in how this technology can be leveraged to improve immersion, engagement, and educational outcomes for CLs. Our research addresses the urgent need for a smart and expressive training tool by developing patient simulators capable of realistically synthesizing non-verbal asymmetric facial cues that are important for the rapid diagnosis of neurological emergencies, such as stroke.

In this paper, we build on our prior work developing patient simulators capable of expressing pain [12, 13] and Bell's Palsy [14], and introduce patient simulators capable of conveying stroke. To our knowledge, expressive patient simulators have not been explored in this way, so this work represents a new contribution to the field.

Our work sits at the intersection of HRI, computer vision, and clinical education, to enable socially interactive robots and virtual agents to simulate human-patient-like FEs and interaction with CLs. We introduce a new control framework called the analysis-modeling-synthesis (AMS) framework. The AMS framework consists of three main components: Facial Expression Analysis (FEA), Facial Action Modeling (FAM), and Facial Synthesis and Animation (FSA). The FEA component enables the AMS framework to detect and track FEs in real-time. The FAM component generates models of a person's face patterns and the temporal aspects of facial deformation. This component also overlay the Facial Paralysis Mask (FPM) derived from a computational model of patient-driven FEs onto the generated models. Finally, the FSA component is responsible for automatically synthesizing tracked facial movements onto robotic and/or virtual faces, and animating their facial components. Using our AMS framework, we are able to discover platform-independent methods to control the FEs of both robots and virtual agents.

Methods

We developed techniques to automatically model the facial characteristics of stroke in real-time, and use the model to synthesize stroke on RPS, ARS, and VPS systems.

Data collection The performance of the stroke modeling system is closely related to the quality of the recorded source videos, which itself is dependent on a number of parameters including patient's background color, distance from the camera, brightness, etc. We are also cognizant of bias in model building, both regarding the representation of different faces, as well as assumptions incorporated into models by researchers [21]. Thus, we sought to follow best practices [10, 1] in ensuring our initial data calibration methods were as inclusive as possible with regards to race and gender, particularly with regard to our intended future population of participants. Therefore, we ran a baseline experiment to determine the setup for achieving the best recording results.

Using this setup, we plan to record source videos from

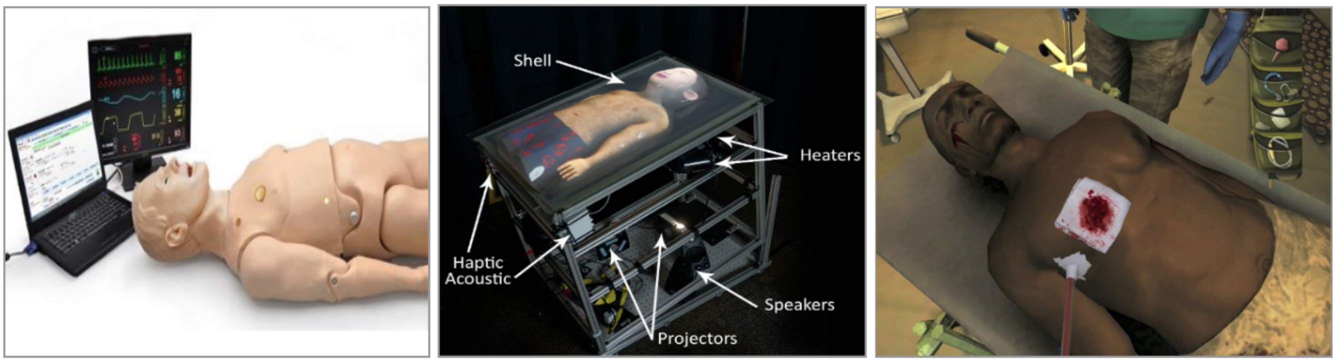

Figure 2: Examples of three types of patient simulators used in simulation-based clinical learning (SBL). Left: , Code Blue III by Gaumard Scientific [6], an example of robotic patient simulators (RPS). Center: CliniSpace [2], an example of augmented reality simulators (ARS). Right: ARS with rare-projected imagery (RPI) [4], an example of virtual patient simulators (VPS)

30 patients' faces admitted to an urban hospital's neurocritical care facility, while clinicians perform a number of selected neurological assessments during examinations [17]. The targeted patients include individuals who have experienced acute stroke resulting in neurological symptoms such as facial droop, eyelid apraxia, dysarthria, and coma.

Detecting and Modeling Stroke In order to build robots and virtual avatars that can replicate realistic, understandable, human-like expressions, it is necessary to be able to detect and model FEs in people. For this reason, we introduced the FPM framework, which consists of two parts: 1) A detection system to extract facial feature points (FFs) from the collected videos using shape-based modeling techniques, and 2) A modeling system to build computational models representing the facial characteristics of stroke in real-time.

For the detection system, we employed an automatic facial action coding system (FACS) system that has three main stages. First, pre-processing, which consists of noise reduction, face region detection, FE detection, data augmentation, and face normalization.

Second is feature extraction and selection, which extracts FFs from source videos of patients using a Constrained Local Model-based approach (CLM) [3]. CLM locates the location of facial landmarks in an input image based on the global statistical shape models and the local appearance information around each landmark. This process yields anonymous features from patients' faces which characterize the clinical characteristics of stroke.

Third is facial feature classification, where a Convolutional Neural Network (CNN) classifier categorizes the FFs to predict specific expressions. These expressions include the FFs tested during neurological assessments to determine impairment on patients with symptoms of stroke.

The structure of the network is as follows. An input layer receives an RGB frame image. The input passes through three convolutional layers, each layer consists of a filter layer and a max-pooling layer. The network has a fully connected layer with 512 hidden neurons and a softmax output layer for classification. We applied a dropout of 0.25 after the last convolutional layer to reduce over-fitting. For training our network, we applied methods of data augmentation, which made small changes to the image to enlarge the training dataset and increase the generalization of our models. For this purpose, we applied a combination of random transformations including vertical flips, outward or inward scales, cropping, and Gaussian Noise.

For the modeling system, we analyzed the visual and physical effects caused by the movements of FFs, and built a novel data-driven model which replicates both the face model and deformations model based on real data. The face model represents patterns that models the human face, both in its neutral state and in states representing the changes in facial movements to display different expressions (e.g, shape, appearance). The deformations model represents patterns of the temporal aspects of facial deformation (e.g., acceleration, peak, amplitude). This technique makes it possible to create FPM models: accurate computational models representing the facial characteristics of stroke.

Synthesis Stroke on Patient Simulators We then employ our AMS framework to track a face in a live facial video stream, overlay prebuilt FPM models onto the stream to generate dynamic asymmetric expressions, and then transfer the expressions onto the face of a robot or virtual agent. Our AMS framework provides the facial movement vocabulary that maps the developed model of facial movements and densities onto the mesh topology of the social robot or virtual avatar's head to make it able to display facial movements corresponding to developed facial expression models.

Considering the diverse clinical backgrounds of patients in the real world, we plan to generate approximately 12 ARS systems which convey different genders, races, and ages.

Planned Evaluation

We will verify the performance of our systems and evaluate our patient simulators using both computational and perceptual methods [14] to identify the most reliable representations of acute neurological injury such as stroke, brain death, and coma. We will further investigate the efficacy of multiple machine learning and deep learning algorithms for rec-

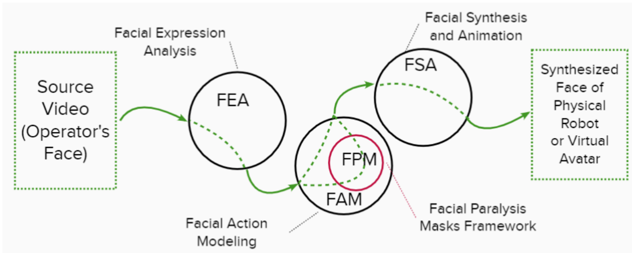

Figure 3: The data flow of the proposed Analysis-Modeling-Synthesis (AMS) framework.

ognizing and modeling expressions that mimic people with different pathologies. In addition to our computational evaluation, we will also conduct a qualitative, expert-based experiment to rate the similarity of our stroke patient simulators with real patients, and study their realism.

We plan to study the effect of using our expressive simulators on improving CLs' skills in decoding visual signals of stroke during live clinical simulations. We will also study different interaction strategies while the CLs interact with these expressive robotic and virtual simulators, including educational performance and engagement.

Discussion

Our research using patient-inspired FEs explores the concept of creating an advanced training tool which supports the clinical education community. Our work provides several potential benefits for CLs by making virtual, augmented, and robotic simulators of acute stroke more varied, interactive, and realistic [17]. Our intelligent RPS, ARS, and VPS systems have a wider range of expressivity than commercially available simulators, including the ability to express neurological impairment such as BP and stroke. First, it enables clinicians to learn the neurological symptoms of stroke in patients with diverse backgrounds. Second, these patient simulators provide CLs with a platform to practice the skills without fear of harming real patients. Finally, these systems can help CLs gain self-confidence in their ability to perform the neurological examination.

This work has several impacts on the greater applied AI, HRI, and clinical informatics communities. It enables researchers to explore new automatic methods for synthesizing asymmetric FEs. It also supports improving the diversity of patient simulators, and enables more interactive and immersive experiences CLs. By leveraging the use of expressive patient simulators, CLs will have the potential to more accurately diagnose people with stroke, and be better able to interact and engage with them. This system will also help enhance immersion in simulation by providing CLs with the opportunity to participate in a comprehensive, realistic experience.

References

[1] Cirillo, D.; Catuara-Solarz, S.; Morey, C.; Guney, E.and Subirats, L.; Mellino, S.; Gigante, A.; Valencia, A.; Rementeria, M. J.; Chadha, A. S.; and Mavridis, N. 2020. Sex and gender differences and biases in artificial intelligence for biomedicine and healthcare. *NPJ Digit Medicine*.

[2] CliniSpace. Clinispace offers healthcare training applications and engine platform. https://www.healthysimulation.com/5499/clinispace-offers-healthcare-training-applications-engine-platform/. Accessed: 06-20-2020.

[3] Cristinacce, D., and Cootes, T. 2008. Automatic feature localisation with constrained local models. In *Pattern Recognition*.

[4] Daher, S.; Hochreiter, J.; Norouzi, N.; Gonzalez, L.; Bruder, G.; and Welch, G. 2018. Physical-virtual agents for healthcare simulation. *Int'l Conf. on Intelligent Virtual Agents*.

[5] Donkor, E. S. 2018. Stroke in the 21 st century: A snapshot. *NCBI, Stroke Research and Treatment*.

[6] Gaumard. 2015. Gaumard simulators. http://www.gaumard.com/aboutsims/. Accessed: 06-20-2020.

[7] Goldberg, C. 2020. Ucsd practical guide to clinical medicine.

[8] James, J. T. 2013. A new evidence-based estimate of patient harms associated with hospital... *Patient safety*.

[9] Jeffries, P. R. 2007. Simulation in nursing education: from conceptualization to education. *NLN*.

[10] Jo, E. S., and Gebru, T. 2020. Lessons from archives: strategies for collecting sociocultural data in machine learning. *Conf. on Fairness, Accountability, and Transparency* 306–316.

[11] Kononowicz, A. A.; Woodham, L. A.; Edelbring, S.; Stathakarou, N.; Davies, D.; Saxena, N.; Tudor Car, L.; Carlstedt-Duke, J.; Car, J.; and Zary, N. 2019. Virtual patient simulations in health professions education. *JMIR*.

[12] Moosaei, M.; Das, S. K.; Popa, D. O.; and Riek, L. D. 2017. Using facially expressive robots to calibrate clinical pain perception. *ACM/IEEE Int. Conf. on Human-Robot Interaction*.

[13] Moosaei, M.; Gonzales, M. J.; and Riek, L. D. 2014. Naturalistic pain synthesis for virtual patients. *Int. Conf. on Intelligent Virtual Agents* 295–309.

[14] Moosaei, M.; Pourebadi, M.; and Riek, L. 2019. Modeling and synthesizing idiopathic facial paralysis. *Int'l Conf. on Automatic Face and Gesture Recognition.*

[15] Morton, P. 1995. Creating a laboratory that simulates the critical care environment. *Critical Care Nursing*.

[16] Pourebadi, M., and Riek, L. D. 2018. Expressive robotic patient simulators for clinical education. *ACM/IEEE Int. Conf. on Human Robot Interaction, Robots for Learning workshop*.

[17] Pourebadi, M.; LaBuzetta, J. N.; Gonzalez, C.; Suresh, P.; ; and Riek, L. D. 2020. Mimicking acute stroke findings with a digital avatar. *Int'l Stroke Conf.*

[18] Riek, L. D. 2017. Healthcare robotics. *Communictions of the ACM, Vol. 60, No. 11. pp. 68-78*.

[19] Shrivastava, S. R.; Shrivastava, P. S.; and Ramasamy, J. D. 2013. Reduction in global burden of stroke in underserved areas. *J. of Neur. in Rural Practice*.

[20] Tarnutzer, A. A.; Lee, S.; Robinson, K. A.; Wang, Z.; Edlow, J. A.; and Newman-Toker, D. E. 2017. Ed misdiagnosis of cerebrovascular events in the era of modern neuro imaging. *Am. Acad. of Neuro.*

[21] West, S. M.; Whittaker, M.; and Crawford, K. 2019. Discriminating systems: Gender, race and power in ai. *AI Now Institute*.