

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2017.DOI

# Facial Expression Rendering in Medical Training Simulators: Current Status and Future Directions

Thilina Dulantha Lalitharatne<sup>1</sup>, Yongxuan Tan<sup>1</sup>, Florence Leong<sup>1</sup>, Liang He<sup>1</sup>,  
Nejra Van Zalk<sup>1</sup>, Simon de Lusignan<sup>2</sup>, Fumiya Iida<sup>3</sup> and Thrishantha Nanayakkara<sup>1</sup>

<sup>1</sup>Dyson School of Design Engineering, Imperial College London, London SW7 1AL, UK.

<sup>2</sup>University of Oxford, Oxford OX1 2JD, UK.

<sup>3</sup>University of Cambridge, Cambridge CB2 1TN UK.

Corresponding author: T.D Lalitharatne (e-mail: t.lalitharatne@imperial.ac.uk).

This work was supported by the Robopatient project funded by the EPSRC Grant No EP/T00603X/1.

**ABSTRACT** Recent technological advances in robotic sensing and actuation methods have prompted development of a range of new medical training simulators with multiple feedback modalities. Learning to interpret facial expressions of a patient during medical examinations or procedures has been one of the key focus in medical training. This paper reviews facial expression rendering systems in medical training simulators that have been reported to date. Facial expression rendering approaches in other domains are also summarised to incorporate the knowledge from those works into developing systems for medical training simulators. Classifications and comparisons of medical training simulators with facial expression rendering are presented, and important design features, merits and limitations are outlined. Medical educators, students and developers are identified as the three key stakeholders involved with these systems and their considerations and needs are presented. Virtual-physical (hybrid) approaches provide multimodal feedback, present accurate facial expression rendering, and can simulate patients of different age, gender and ethnicity group; makes it more versatile than virtual and physical systems. The overall findings of this review and proposed future directions are beneficial to researchers interested in initiating or developing such facial expression rendering systems in medical training simulators.

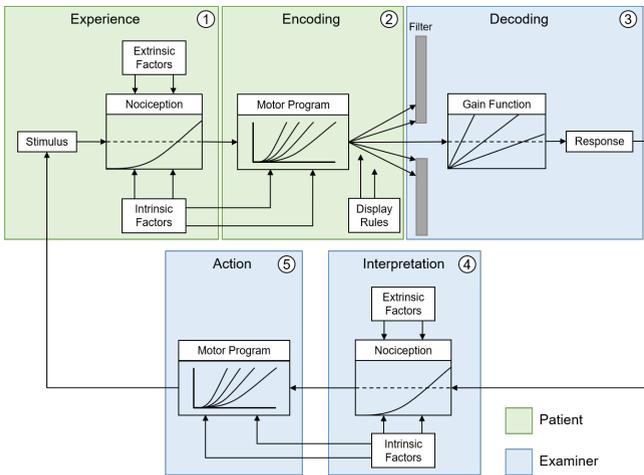
**INDEX TERMS** Facial expressions, Facial expression rendering, Medical simulators, Medical training, Robotic patients, Human-Machine Interaction

## I. INTRODUCTION

**R**ETROSPECTIVE medical record reviews in British hospitals [1] show that approximately 10% of patients admitted to a hospital experience unintended injuries caused by medical management. A report by the Institute of Medicine [2] showed that up to 98,000 hospital deaths occur in the United States of America as a result of medical errors each year. Enhancement of techniques available to train medical professionals to integrate multiple sources of information to make correct conclusions can reduce such accidents. Some physical examination procedures such as palpation [3] require years of experience to acquire the right motor skills, perceptual strategies, and therapeutic attitudes for general practitioners (GPs) [4], [5]. Effective education of practical skills can positively influence clinician behaviour, reducing the risk of patient harm [6], [7]. Traditional teaching

methods, such as live demonstrations followed by students practising under tutor supervision, are often perceived ineffective due to class time being consumed by teaching and demonstration of practical skills, availability of teaching staff, and resource considerations [8].

Medical procedures such as physical examinations involve physician-patient interaction, a process involving multiple feedback modalities, and is affected by the intrinsic and extrinsic factors of both parties. Intrinsic factors such as the patient's pathological conditions and mood states; Extrinsic factors such as gender, age, and ethnicity have been shown to affect how the patient expresses the feeling of pain [10]. Fig. 1 shows an overview of this interaction system, where the interaction points occur between phase 2 and 3, and phase 5 and 1. Physicians can learn some of the extrinsic factors from the patient's appearance and medical history,



**FIGURE 1.** Example of physician-patient interaction system diagram during physical examination (adapted from Prkachin, 1994 [9]).

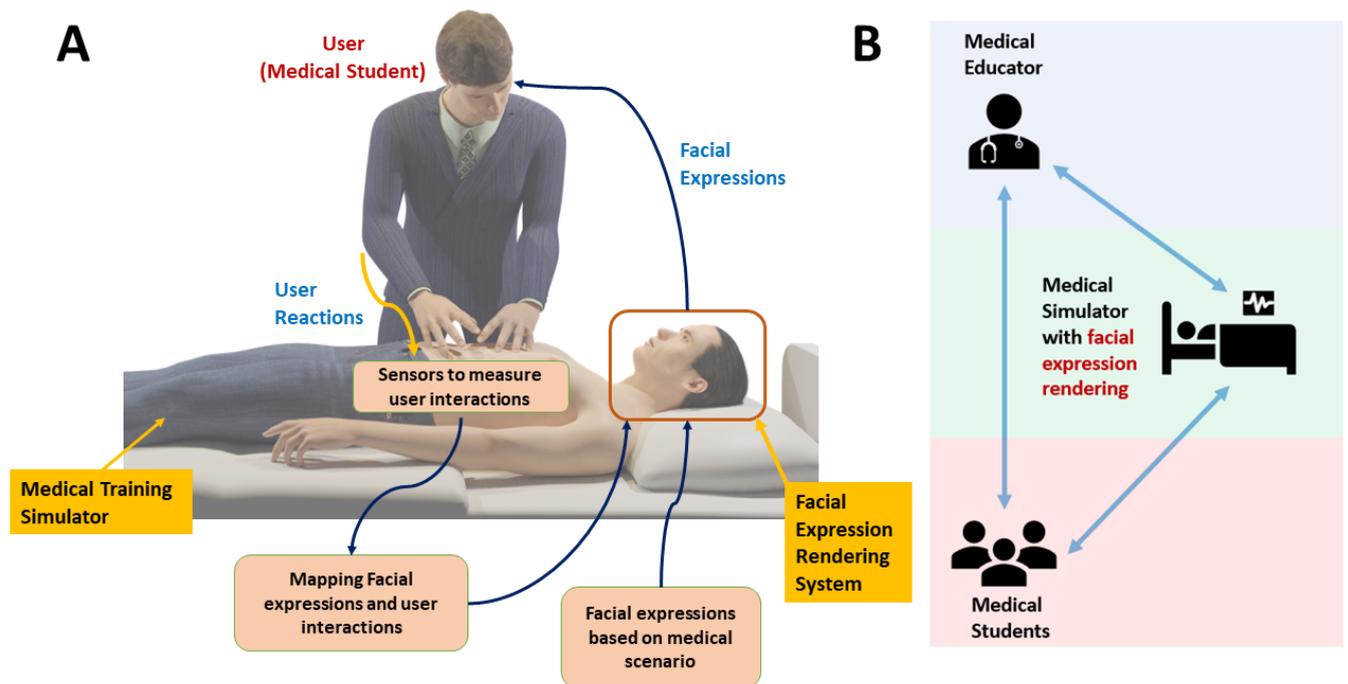
whilst some of the intrinsic factors that directly relate to the patient's current health conditions remain unknown. Physicians attempt to engage with the patient's experience and the encoding process by decoding and interpreting verbal response, and combine with the visual and haptic feedback to determine the unknown factors to make diagnoses.

Visual feedback, including static information conveyed by the patient's general appearance, and dynamic information such as changes in the patient's facial expressions, communicate more information than verbal and vocal messages [11]. Mehrabian [12] found 55% of the total impact of a message is from the non-verbal component of communication, and Pease [13] found non-verbal messages can be conveyed by a person's general appearance, personal hygiene, clothes and etc. Visual perception through facial expressions during primary medical examination such as manual palpation is often used as a feedback modality to assess a range of medical hypotheses for diagnosis. When patients experience discomfort or pain, they communicate their feelings and emotions through verbal and non-verbal channels, including facial expressions [3]. Pain is one of the most commonly exhibited reactions during medical examinations. It can be communicated through facial expressions, which has been the most direct method for humans to communicate their feelings and emotions since infancy [14]. When properly interpreted, information related to the pain experienced by the patient conveyed to the examiner by facial expressions is sometimes more accurate than verbal responses [15]. But if the observer wrongly diagnoses the severity of one's pain, it could result in mistreatment or even mortality [16]. The intrinsic and extrinsic factors in both the patient and the physician increases perceptual subjectivity when perceiving visual feedback. Hence it is important for the students to experience treating patients of different demographics and backgrounds, as misjudgments in visual feedback due to subjectivity could result in distrust between patients and health workers [17], [18].

Simulation-based education (SBE) is frequently used in recent years, as it can provide safe and effective learning environments for students [19]. Its central tenet is student-tailored making and learning from mistakes and errors, which are prevented or immediately terminated to protect the patient in clinical settings [20], [21]. Using SBEs, the causes of mistakes made during the training can be reviewed openly without liabilities, allowing students to confront the mistakes and recognise the importance and value of the experience, and possibly improving the quality of event reporting [20], [22]. Standardized patients (SPs), physical training simulators such as manikins, and computer-based simulation or virtual training simulators [23] are the main SBE types employed by medical educators. They range from low to high fidelity for different medical specialties such as internal medicine, emergency medicine, surgery and dentistry.

SPs are healthy people acting as a patient in a consistent and standardized manner [24]. They are of the highest fidelity as a simulator but are difficult to train, and are unable to exhibit desired physiological conditions and symptoms at their will. Moreover, it is difficult to find SPs to represent some demographic groups such as children and infants [25]. Physical and virtual training simulators do not generate all types of feedback SPs do. Haptic, auditory, and visual feedback are often regarded as the most important modalities when it comes to learning and training of medical operations. Many physical training simulators, such as manikins, are designed to provide accurate haptic feedback to help improve clinicians' motor skills, and many guidelines and design considerations were provided for haptic response and haptic visualisation [26]. Manikin simulators with haptic feedback can be classified by its fidelity, where low-fidelity manikins [27] are used to train specific tasks or procedures, and high fidelity manikins [28] can be used to simulate a greater variety of medical conditions. Although manikin simulators are widely used in medical training and students are more comfortable with using manikin simulators than SPs [29], they cannot be easily customised to simulate patients of different demographic characteristics. On the other hand, systems such as computer-based simulators and virtual patients provide great flexibility in simulating patients of different sex, age, and ethnicity. But they often lack in compatibility with physical sensing and haptic feedback.

As previously emphasised, it is important to render accurate facial expressions in medical training simulators irrespective of the types. An overview of the implementation of a medical training simulator with visual feedback through facial expression rendering is shown in Fig. 2A. This closed-loop process has two main interaction points: facial expressions generated by the simulator and perceived by the student, and the subsequent reaction generated by the student and exerted on the simulator. The facial expression rendering system can encode inputs from the sensors embedded in the simulator and display real-time dynamic facial responses. Fig. 2B shows the stakeholders involved in this system. Researchers and developers should consider both the educator



**FIGURE 2.** (A) A conceptual overview of a typical medical training simulator with facial expression rendering system. (B) Main stakeholders (i.e: medical educators, students and simulator developers) involved in such system.

and student needs when designing the system.

The primary focus of this paper is to review facial expression rendering techniques used in medical training simulators that have been reported to date, and to identify important design features, merits and limitations of these systems. The literature review on facial expression rendering in medical training simulators was performed based on search terms medical training simulators with pain expressions, eye movements, and facial expressions using Google and Google Scholar, and both commercial products and research projects were included. Alternative versions of the same system were excluded, and the systems that sorted out were classified into physical, virtual and hybrid systems. Even though there are many reviews of SBEs [19], [30] and medical training simulators with feedback methods such as haptic feedback [31], [32], to the best of our knowledge, no attempts have yet been reported on in-depth reviews of facial expression rendering approaches in medical training simulators. A timely written review paper on facial expression rendering in medical training simulators may help, not only to identify the current status of the research, but also to provide information to future researchers interested in developing such systems.

The paper is organized as follows: Section II presents an overview of facial expressions, specifically pain expressions, as it is one of the most common and useful feedback modalities of the patient during physical examinations, and facial expression models. Section III reviews existing facial expression rendering systems in medical training simulators

with classifications and comparisons. Section IV presents a brief review of facial expression rendering approaches in other research domains to explore the potential to incorporate the knowledge from these works into developing facial expression rendering systems for medical training simulators. Sections V and VI summarise the findings of this review, and concludes the paper with a discussion of design considerations for facial expression rendering systems in medical training simulators and potential future directions.

## II. FACIAL EXPRESSIONS

Many of the practical skills clinical students need to acquire are based on sensory-motor coordination, which consists in generating, accessing and evaluating useful sensory information through actions. The key difference in diagnostic skills between an experienced doctor and a student is the ability to focus, identify and distil perceptual and haptic information that can justify hypotheses in a short time. Therefore, medical simulators should prompt students to use active perception in conjunction with information gain from haptic feedback, and this can be done by incorporating visual feedback into the system. Visual feedback via facial expressions is an important source of information in medical examination. It is one of the most direct methods for the patient to communicate emotions and feelings, sometimes involuntarily [15]. Needless to say, understanding facial expressions is particularly important when the patient is unable to utilise verbal communication [33]. However, the challenge is to train a

physician to extract examination-related information out of a complex repertoire of expressions with subjectivity and biases. In this section, we will explore theoretical frameworks of facial expression generation, common features of facial expression across ethnic groups, the facial expressions of pain, and facial action coding systems.

### A. MECHANISMS OF FACIAL EXPRESSION GENERATION

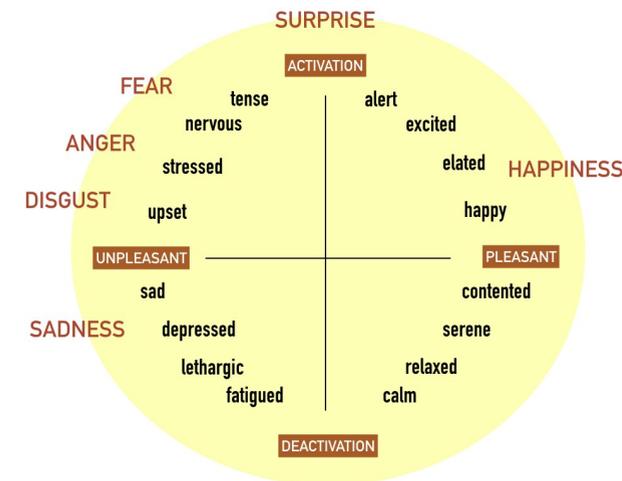


FIGURE 3. Map of core affect (adapted from Russell and Barrett, 1999 [34]).

Emotions are typically conveyed via facial expressions, voice intonation, gestures, postures and behaviour [35]. In everyday life, people differentiate between a variety of emotional episodes, such as resentment, guilt, sadness, anger, or happiness [36], and rely on the idea that faces reveal what others feel inside (i.e., emotional expressions), as well as what information is inferred from facial expressions (i.e., emotion perception) [37]. The vast majority of the literature in psychology and neuroscience has relied on the common or classical view of emotions, assuming they have distinct boundaries and neural pathways observed in the brain and body representing reactions to behavioural or contextual stimuli [38]. For example, emotions are presumed to be “read” via prototypical facial configurations representing discrete universal categories such as anger, fear, disgust, happiness, sadness and surprise [37], [39], [40]. Researchers often use Facial Action Coding Systems or FACS, which are schemes that allow for combining various parts of the face into emotional prototypes, where participants are asked what the person on the picture is feeling. For example, the pain expression may include brow lowering, skin drawn tightly around the eyes, and a horizontally stretched open mouth with deepening of the nasolabial furrow, and anger includes furrowed brows, wide eyes, and tightened lips [41], [42].

Nonetheless, assumptions about configurations of emotional expressions have not been reliably replicated across samples, cultures, nor situations/contexts [37]. In terms of other measurement modalities, two recent large meta analy-

ses on emotions and brain locations indicate large discrepancies between individuals, as emotions such as disgust, sadness, anger, fear and happiness cannot be pinpointed to activity in specific regions of the brain using classical tasks of measurement [43], [44]. In addition, the experience of emotional events (a common way to identify emotional states) does not take into account individuals who experience emotional events without direct awareness, nor those who struggle with words to express and describe their own and others’ emotions, such as alexithymics [37]. Despite an exponential increase in applications modelling human emotions in areas such as affective computing, these models provide low predictive capability, often resulting in inaccurate emotion detection in computerised systems [45], [46].

The constructionist view is an opposing framework where emotions represent dimensions rather than categories. Commonplace mental states including perceptions, cognitions, and emotions are assumed to be constructed from a combination of fundamental psychological operations [47]. Viewed as so-called core affect, emotions can be described as a linear combination of two underlying, largely independent neurophysiological systems known as valence and arousal (depicted in a circumplex Cartesian space in Fig. 3, where the inner circle shows a schematic map of core affect and the outer circle shows where several prototypical emotional episodes typically fall). The valence system determines the degree to which an emotion is pleasant or unpleasant, and the arousal system determines the degree to which it is behaviourally activating. Consequently, a 90 degrees distance (e.g., “happy” and “tense”) represents independent, whereas a 180 degrees distance represents opposite affective experiences (e.g., “happy” and “sad”). Importantly, experiences of emotions are considered to be ambiguous, overlapping sensations that are the product of activity in neural pathways subserving valence and arousal, but which become contextualized and classified through the processing of relevant situational, historical, behavioural, and physiological cues that people use to safely navigate the world [48].

Thus, rather than being ‘caused’ by external or internal events, emotional experiences are instead predicted based on information from core affect and context [49]. In this view, the brain is a probabilistic prediction processor or engine [50], and core affect can be used to predict the precision (or granularity) of emotional experiences [48]. One illustration of this is that viewing Serena Williams celebrating a Wimbledon victory allows viewers to identify her joy (a form of happiness) only after viewing the picture with a tennis racket in full gear (i.e., with contextual information indicating a win). Evaluating a close-up of her face without knowing the context would produce an entirely different emotional evaluation, such as that of rage [46]. This ambiguity is also seen in people with mental health problems and in chronic pain [51], [52]. Thus, emotions should be modelled as likelihoods of emotional reality rather than reactions to environmental stimuli. This has direct implications for computational modelling forming the basis for development of

accurate emotion mapping of pain, as lack of universality in emotional expression as well as the context of medical examination need to be taken into account.

### B. FACTORS UNDERPINNING THE DIVERSITY OF FACIAL EXPRESSIONS

Current literature in psychology and neuroscience show large inconsistencies in findings on the foundation as well as universality of human emotions, with growing work indicating that people are not always good at recognising emotions from faces without prior contextual clues [49]. In addition, problems with reliability, specificity, generalizability, and validity are largely present in studies based on the view that emotions are categorical [37].

For instance, a seminal study by Ekman [53] identified commonality in different cultures for six displayed emotions: happiness, sadness, fear, disgust, surprise and anger; suggesting understanding one's emotion and feeling by observing facial expression is a reliable method when there are barriers in verbal communication. Izard [54] supported this idea of universality in facial expression and took inspiration from Tomkin [55] who proposed the Discrete Emotion Theory, stating that there is a limited number of pancultural basic emotions.

Nevertheless, the authority of this work has been questioned, as subsequent studies removing clues about context have found less congruence across cultures in emotion perception from facial information [37]. For instance, there are differences in interpretations of certain emotional expressions. Despite assuming that emotions such as anger will be easily identifiable by participants using behavioural and/or physiological measures (e.g., via facial or auditory expressions, or brain scans), there is low correspondence across such measures to reliably indicate a universal instance of emotions in the brain or body [38]. Comparing East Asian and Western cultures, Chen *et al.* [56] found differences in perceptual discrimination of facial expression of pain and orgasm based on facial actions. Such inconsistencies are also obvious in recent attempts at technological applications (e.g., Microsoft's Emotion API), with no emotion recognition algorithm thus far reliably predicting either people's emotion expressions or perception from facial information alone [37].

In sum, facial expressions can be viewed as a communication medium with embedded subjectivity in the observer's perception caused by differences in demographic attributes such as gender, age, and ethnicity [57]–[59]. In medical training, therefore, the trainees should be exposed to patients of a great variety of demographic characteristics to increase perceptual objectivity.

### C. PAIN EXPRESSIONS

Pain is defined as "an abstract concept which refers to (1) a personal, private sensation of hurt; (2) a harmful stimulus which signals current or impending tissue damage; (3) a pattern of responses which operate to protect the organism from harm" [60]. Clinicians and doctors rely on pain as a feedback

modality to adjust their examination procedure accordingly to minimise the discomfort the patient experiences [61].

Pain expressions have been identified as common and useful feedback to the clinician. They are subjective by nature and the intensity of pain have been proven to be more accurately measured by analysing facial expressions than self-reports of the patient [15], [61], [62]. It is crucial to understand how pain is generated, expressed, observed and understood. If the observer wrongly diagnoses the severity of one's pain, it could result in mistreatment or even mortality [16]. Pain intensity can be measured using the verbal rating scale (VRS-4), the visual analog scale (VAS) and the numeric rating scale (NRS-11). Breivik [63] showed VRS-4 was less sensitive than VAS and NRS-11, and VAS showed similar sensitivity in relation to the pain intensity, highlighting the importance of analysing facial expressions in evaluating pain. There are also pain scales for assessing children's pain such as the Faces Pain Scale-Revised [64], and the Wong-Baker FACES Pain Rating Scale [65]. These scales are based on facial expressions as children are less able to express their feelings verbally.

Pain can be communicated and interpreted via nonverbal communication methods. Rosenthal *et al.* [66] presented a nonverbal communication model as a three step process: A: internal state of an experience. B: encoding A into expressive behaviours. C: observers draw inferences about the sender's experience. Ekman's Neurocultural Theory [67] suggested that the facial affect program (one-to-one map between felt emotion and displayed facial expression) is the same for all people in all cultures. But people use "management techniques" to control and sometimes override the operation of the universal facial affect program under some social settings. Prkachin included the social stimuli in Rothenthal's model and proposed a general model of a pain episode [9], which is composed of three stages: Experience, Encoding and Decoding. This model outlined the variability of the relationship between the intensity and quality of pain, and one's expression of pain.

Many researchers evaluated how pain is translated to facial expressions. Chapman and Jones [68] evaluated reactions to pain stimulus with heat. They found similarities between peoples' facial expressions such as the contraction of the eyelids, and noted that the majority of the subjects were unable to constrain that muscle movement when asked to. Prkachin [16] found similarities in facial features when people experience pain.

Therefore, medical training simulators should be able to render the facial expressions of pain to a high level of similarity of the human expression to stimulate the visual information integration process of the trainee. This can be achieved by synthesising specific facial features and muscle groups. The next section introduces systems and frameworks for describing facial expressions.

TABLE 1. Facial Expression Evaluation Methods

Reference	Use FACS	Methodology and Remarks
Suwa et al., 1978 [69]	No	One of the first studies on automatic facial expression analysis.
Terzopoulos, 1990 [70]	No	Used anatomical models of the facial muscle actuators to identify expressions, enabling accurate synthesis of transient expression by computing muscle actuator controls.
Mase et al., 1991 [71]	No	Used optical flow to extract muscle actions, and categorised the image sequences into classes of facial expression using a feature vector of time and local spatial area of the variance.
Neely et al., 1992 [72]	No	A recognition method focused on facial movements of forehead wrinkling, eye closing, nose wrinkling, and mouth smiling, by computing intensity of pixels in the predefined regions from subtracting frames of the facial motion from a reference image.
Wood, 1994 [73]	No	Used video microscaling to determine distances in facial movements by superimposing a computer-generated measuring scale over a video recorded image of facial movements.
Yacoob and Davis, 1996 [74]	No	Used optical flow in model building, with a focus on identifying the direction of motions caused by facial expressions.
Curtis and Garrison, 1996 [75]	No	A emotion recognition system using feed-forward neural networks to detect six basic emotions by analysing static images.
Roenblum and Yacoob, 1996 [76]	Yes	High success rate in correctly identifying facial expressions using a radial basis function network.
Yuen et al., 1997 [77]	No	Used Moiré topography to evaluate the degree of facial palsy as an objective method in addition to other subjective macroscopic evaluation methods.
Essa et al., 1997 [78] and Fukui et al., 1998 [79]	Yes	Used shape matching and pattern recognition. Facial motions are encoded with geometric models and that distinctive features such as pupils, nostrils, and mouth edges are extracted.
Cohn et al., 1999 [80]	Yes	Quantitatively measured image sequences and encoding facial expressions into FACS AUs.
Donato, 1999 [81]	Yes	Used optical flow, principal component analysis and Gabor wavelet transform for automatic facial action recognition with FACS using image sequences.
Zhang, 1999 [82]	Yes	A two-layer approach to capture geometric positions of fiducial points and Gabor wavelet coefficients at those points. Both computational efficiency and generalisation performance could be improved by discarding points that carry little information such as the ones on cheeks and on forehead.
Lien et al., 2000 [83]	Yes	Used feature extraction and FACS classification. Accurately decoded facial expressions of subjects of varied ethnicity.

#### D. FACIAL EXPRESSION MODELS

As previously mentioned in Section II-A, FACS was one of the first framework to systematically decode facial expressions. FACS divides common facial expressions into smaller movement primitives known as action units (AUs), allowing facial expressions to be analysed quantitatively. The Emotional Facial Action Coding System (EMFACS) was later developed to improve computational efficiency of FACS. There are also other facial action coding systems. Izard created the Maximally Discriminative Facial Movement Coding System (MAX) [84], which is commonly used to examine infants' facial expressions. The System for Identifying Affect Expressions (AFFEX) [85] is another facial action coding system for children, which makes direct judgements of the affective meaning of the patterns of facial muscle movements in infants and young children. Ekman's system was more quantitative and it was later widely used in automatic facial expression recognition systems, whereas Izard's was more qualitative and was more commonly used in medical training.

Many models and systems for analysing and generating facial expressions were developed using the facial action coding frameworks and data analysis and image processing techniques. Deception means to retrieve information from one's expression [15], and it can be classified into two categories: subjective and objective [86]. Objective methods rely on taking measurements and evaluating numerical differences to determine changes and states of various facial expressions. Many objective methods were developed using

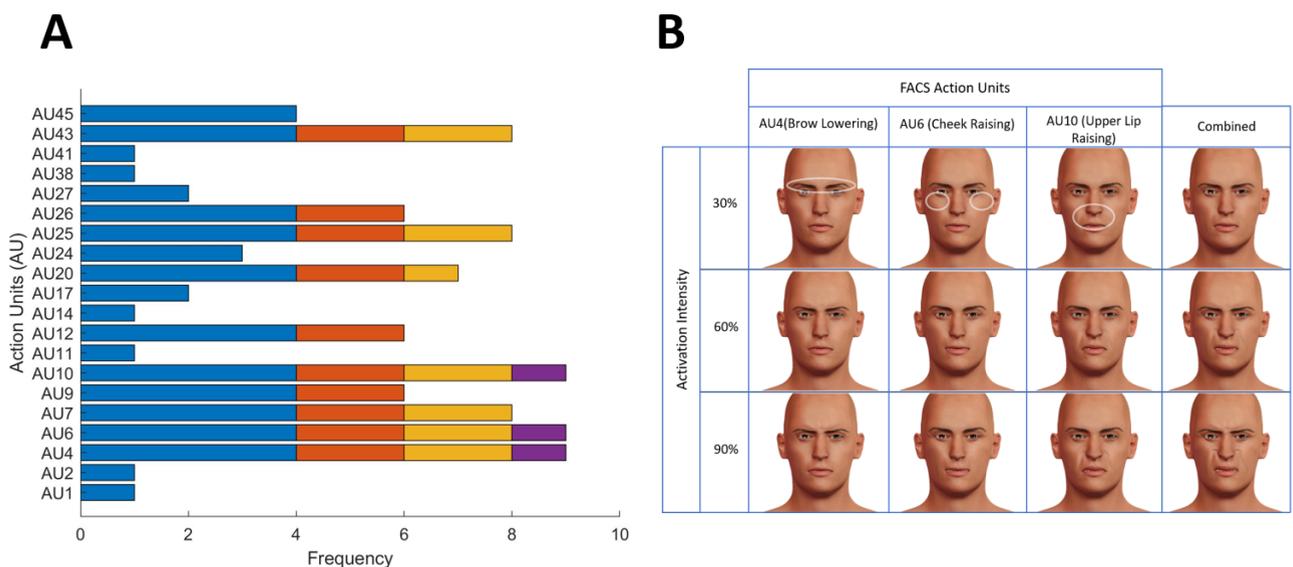
signal processing and data analysis methods as summarised in Table 1. Subjective methods correspond to various facial nerve grading systems, and the most common system is the House-Brackmann facial nerve grading system [87]. Subjective methods are esteemed to be less reliable than objective methods due to variations caused by observer bias, and is not commonly used in facial expression decoding systems.

Researchers found the pain expression could be identified as a combination of at least four AUs using FACS. Patrick [88] used FACS to evaluate peoples' responses to a broad range of pain stimuli and found significant increase in 4 AUs with pain intensity across all categories of pain: brow lowering, orbit tightening, upper-lip raising/nose wrinkling, and eye closure. Craig and Patrick [89] conducted what is known as the cold pressor test and found six AUs associated with pain. Craig [90] analysed all the AUs that activate during pain experience and found lower facial actions more likely to occur during severe pain. Table 2 summarises some of the significant work and findings on pain stimulus, intensity and facial expression up to date.

Fig. 4A shows that some AUs are more frequently shown in pain expressions than others, the colours classify AUs based on their appearing frequency in previous studies. The result shows AU4, 6 and 10 are most likely to be present in pain expression, and this finding is comparable to findings by other researchers. Prkachin [95] found using AU4, 6, 7, 9, and 10 are sufficient to exhibit pain expression at varying intensity levels. Chen et al. [56] found AU4, 6, 9, 10, and 20

**TABLE 2.** Pain Intensity Evaluation using FACS

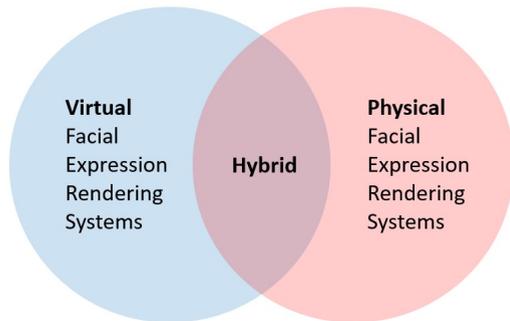
Reference	Experimental Method	Related AUs	Remarks
Craig & Patrick, 1985 [89]	Cold pressor test.	AU7, AU10, AU12, AU25, AU26 and AU43 or AU45	Social models had potent impact on verbal report and pain tolerance but not on facial expression.
LeResche, 1982 [41]	Examine candid images of acute pain.	AU4, AU6, AU11, AU17, AU20, AU25, AU26, AU43	High degree of regularity was shown in facial movements associated with intense, acute pain. Acute, severe pain in adults is characterised by a pattern of facial behavior distinct from patterns shown in disgust, anger, fear or sadness.
Patrick et al., 1986 [91]	Electric Shock.	AU4, AU6, AU10, AU45	Nonverbal expression yields information that is non-redundant with self-report.
Craig et al., 1991 [90]	Examine patients with real, suppressed and fake chronic low back pain.	AU4, AU6, AU7, AU10, AU25, AU43	AU4, 6 and 12 were displayed more frequently in faking pain. Attempt to suppress pain was not successful as residual facial activity persisted.
LeResche and Dworkin, 1988 [92]	Compare FACS analysis with self-report in TMD patients.	AU4, AU6, AU7, AU9, AU10, AU20, AU25, AU26, AU43, AU45	Facial expressions can reliably measure pain and is not disturbed by psychological distress.
Prkachin and Mercer, 1989 [93]	Analyse pain intensity of patients with gleno-humeral joint pain using FACS.	AU4, AU6, AU7, AU9, AU10, AU12, AU20, AU25, AU43	Facial actions are consistent with sensory and affective pain scales. AU4,6, 9 and 12 showed correlation with subjective pain ratings.
Prkachin, 1992 [88]	Expose subjects to electric shock, cold, pressure and muscle ischemia. Measure facial expressions with FACS.	AU1, AU2, AU4, AU6, AU7, AU9, AU10, AU12, AU14, AU17, AU20, AU24, AU25, AU26, AU27, AU38, AU41, AU43, AU45	AU4, 6, 7, 9, 10 and 12 showed increase in activation with pain intensity.
Simon et al., 2008 [94]	Compare pain and basic emotions using FACS.	AU4, AU6, AU7, AU9, AU10, AU12, AU20, AU24, AU25, AU26, AU27	Facial displays of pain are distinctive from basic emotions.
Prkachin, 2008 [95]	Examine facial expression of pain in patients with shoulder pain using FACS.	AU4, AU6, AU7, AU9, AU10, AU12, AU20, AU24, AU26, AU43	AU4, 6, 7, 9, 10 and 43 consisted of a core expression of pain with unitary properties.
Lucey et al, 2011 [96]	Use FACS to code the UNBC-McMaster shoulder pain expression archive database.	AU4, AU6, AU7, AU9, AU10, AU12, AU20, AU25, AU26, AU43	A large FACS coded pain expression database with self-report and observer measures.



**FIGURE 4.** (A) Action Units (AU) employed in studies (n=10) presented in Table 2. AU4, AU6 and AU10 were seen as the frequently used AUs to represent the pain facial expressions (B) Rendered facial expressions of pain with different AU activation intensity. Facial muscle groups driven by the assigned AUs are highlighted by the ellipses.

to be the most frequent AUs for pain expression in Western and East Asian cultures. This plot also suggest the simplest pain expression model can be built by synthesising AU4, 6, and 10, as shown in Fig. 4B. By using different numbers of AUs, researchers and developers can build systems of different degrees of realism, though the individual activation intensity and activation orders should be experimented and tested for dynamic simulations.

### III. MEDICAL TRAINING SIMULATORS WITH FACIAL EXPRESSION RENDERING



**FIGURE 5.** Classification of the facial expression rendering systems based on methods of implementation

The facial expression rendering systems designed for medical training simulators range from fully virtual to fully physical, as classified in Fig. 5. Virtual systems rely on mediums such as computer monitors, televisions or projections on flat screens to render the facial expressions. Physical systems can be active manikins or robotic heads/faces that mechanically render facial expressions. And hybrid facial expression rendering systems are consisted of both virtual and physical components to synthesize the facial expressions. Thus far, different approaches or solutions have been proposed for rendering facial expressions in medical training simulators, as summarised in Table 3.

#### A. PHYSICAL FACIAL EXPRESSION RENDERING SYSTEMS

Full body manikins can simulate many different symptoms. SimMan 3G [28] from Laerdal Medical, an adult patient simulator for advanced training in emergency care, and Pediatric HAL [97] from Gaumard Scientific, an pediatric patient simulator are established commercial manikin simulators. Both simulators render facial expressions mechanically, and can simulate a variety of neurological and physiological symptoms. HAL can simulate emotions through dynamic facial expressions, movement and speech, whereas SimMan 3G is limited to eye movements only. Although only eye movements are implemented as SimMan's facial expression feedback, the state of the eye is a critical factor in recognition of consciousness in emergency situations. SimMan can be customised using wigs to simulate patients of different gender but the skin colour cannot be changed. HAL can be made

in three skin colours: light, medium and dark, but the ethnic features of the face do not differ.

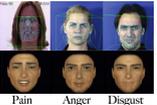
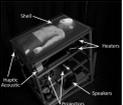
Dentistry is another area where simulators are used to improve motor skills and hand-eye coordination of the students, which are essential clinical skills [107]. There are two main dental simulation systems: SIMROID [98] and DENTAROID [99], which are designed by Japanese researchers with physical facial expression rendering. SIMROID can simulate stress or discomfort the patient feels during the procedure. The face of SIMROID is actuated by air cylinders to physically display facial expressions of discomfort, pain and positive acknowledgements. Its eyes can rotate about the vertical and horizontal axis, and the eye lids can exhibit different degrees of opening. Its mouth can open and close, and its head can rotate about the vertical and horizontal axis. DENTAROID has automatic dialogue patterns, which gives students a more realistic clinical training experience. It helps students to learn to avoid accidents and improve communication competency with patients. DENTAROID is capable of performing different reaction movements that simulate accidents which can occur during treatment, such as reaction to pain. It can also perform facial movements such as eye blinking to make the simulation more realistic. Both robots have skin textures closely resembles the human skin, and their ethnicity and gender cannot be changed or customised.

In addition to the commercial products, many platforms and systems are being explored by different research groups. Baldrighi et al. [100] proposed a bio-inspired facial expressions rendering system for a medical manikin. Their system was able to synthesise 7 emotions including happy, sad, angry, surprised, suspicious, disgusted and sarcastic expressions. The skin of the face was made of silicone elastomer to replicate the texture and appearance of real human skin, and the underlying structure was consisted of 13 DOFs driven by motors. The system was able simulate eyes movements such as blinking, iris dilation, and object tracking. The neck of the manikin was realized with a 6 DOF Stewart platform.

#### B. VIRTUAL FACIAL EXPRESSION RENDERING SYSTEMS

Another approach of facial expression rendering is virtual simulations. In 2014, Moosaei et al. [102] reported an approach for synthesizing naturalistic pain on virtual patients. They used the UNBC-McMaster pain archive and MMI database for acquiring source videos of anger and disgust. A Constrained Local Model (CLM) based face tracker was used to extract 68 feature points frame-by-frame from each source video. To render three different expressions: pain, anger and disgust, they mapped the extracted feature points to control points of an animated virtual character using Steam Source SDK. The authors also claimed that the naturally driven pain synthesis they used in the study was comparable to FACS based pain synthesis. In 2017, Moosaei et al. [106] performed a study on using facially expressive robots to calibrate clinical pain perception. In this study, they developed a virtual patient to synthesise facial expressions of pain, anger and

**TABLE 3.** Medical Training Simulators with Facial Expression Rendering

Product / Study	Application Type	Type	Commercial Availability	Synthesized Facial Expressions	Additional Feedback Capabilities	Diversity Options
 <p>SimMan 3G [28]</p>	Patient simulator	Physical	Yes	Slow, normal, fast blinks and winks, Partially open papillary accommodation.	Variable lung compliance, airway resistance, breathing patterns, patient voice playback.	One body shape with a beige skin tone. Can be customised using wigs.
 <p>Pediatric HAL [97]</p>	Pediatric robotic patient	Physical	Yes	Facial expressions such as Anger, Transient Pain, Ongoing Pain, Amazed, Quizzical, Worried, Anxious, Crying, Yawning and advanced eye movements.	Heart, lung and bowel sounds, breathing patterns.	One body shape with light, medium and dark skin tones.
 <p>SIMROID [98]</p>	Dental robotic patient simulator	Physical	Yes	Discomfort or pain, and positive acknowledgement through facial expressions.	Hand, leg movements, and speech.	Asian female avatar only.
 <p>DENTAROID [99]</p>	Dental robotic patient simulator	Physical	Yes	Reaction to pain, cough reflex, vomiting reflex.	Hand movement and speech.	Asian female avatar only.
 <p>Baldrigi <i>et al.</i>, 2014 [100]</p>	Medical manikin	Physical	No	7 emotions including happy, sad, angry, surprised, suspicious, disgusted and sarcastic. Simulate papillary reaction.	Neck movement and object tracking with the eyes.	Pale silicone elastomer skin with a black wig.
 <p>Wandner <i>et al.</i>, 2010 [101]</p>	Virtual 3D human avatar	Virtual	No	Pain expression for various intensities.	No additional feedback	Variable sex (male, female), race (Caucasian, African American) and age (young adult, older adult).
 <p>Moosaei <i>et al.</i>, 2014 [102]</p>	Virtual patient	3D Virtual	No	Pain, Anger and Disgust.	No additional feedback	Variable sex (male, female)
 <p>Maicher <i>et al.</i>, 2017 [103]</p>	Virtual Patient	Virtual	No	Happiness, Anger, Sadness, Fear, Surprise, Disgust, Contempt, Pain and Character diversity.	Speech-based conversation and virtual gestures.	Not reported
 <p>Kotranza <i>et al.</i>, 2008 [104]</p>	Mixed reality human for breast examination	Hybrid	No	Pain and discomfort.	Speech-based conversation, virtual gestures, facial and verbal feedback based on verbal input and sensor data.	Variable race.
 <p>Rivera-Gutierrez <i>et al.</i>, 2012 [105]</p>	Shader lamps virtual patients	Hybrid	No	Eye movements.	No additional feedback	Not reported.
 <p>Daher <i>et al.</i>, 2020 [106]</p>	Physical-virtual patient simulator	Hybrid	No	Facial expressions could be adjusted to fit different scenarios.	Speech, visual head, arms, and finger movement	Not reported.

disgust, and compared it with the same facial expression rendered on a Philip K. Dick (PKD) humanoid robot. The study found that it was easier for observers to detect pain in a virtual patient than in the high-fidelity, facially expressive PKD humanoid robot patient simulator. In addition, their approach allowed accurate generation of virtual human avatars of various demographic features.

In 2017, Maicher et al. [103] developed a conversational virtual standardized patient to help students practice history-taking skills. Character models in the simulator were created with Autodesk Character Generator and refined using Autodesk Maya. The virtual characters were able to synthesise various emotions via facial expressions such as happiness, anger, sadness, fear, surprise, disgust, contempt and pain. This system also allowed skin color and the gender of the character to be changed during sessions.

Wandner et al. [101] used virtual human (VH) technology with the People Putty software program to generate human avatars of different demographic attributes with various pain intensity expressions. The benefits of using virtual avatars are that the facial features and pain expressions can be standardised without biases from the construction of the stimuli, and that patients of different sex, race, age and pain intensity can be generated rapidly. Using this novel method they found the characteristics of the VHS influenced the ratings of pain assessment and treatment recommendations. This finding suggests that medical training systems should give students exposure to patients with different demographic attributes to minimise the effect of biases in treating real patients.

### C. HYBRID FACIAL EXPRESSION RENDERING SYSTEMS

Facial expression rendering methods are not only limited to virtual and physical approaches. In 2008, Kotranza et al. [104] proposed a mixed reality human platform for breast cancer examinations which we classify as a hybrid approach. Mixed reality or hybrid systems preserve features of both the physical and virtual approaches, often done by projecting virtual models onto physical objects. The physical systems consisted of a full-body physical embodiment, in the form of a plastic manikin. The left breast of the manikin was a soft phantom simulating the feel of breast skin, tissue, underlying breast masses, and contains twelve pressure sensors to detect the user's touch. The virtual system was realized using a head mounted display (HMD) and wireless microphone. Head orientations and positions of the user were tracked using a combination of infrared-marker based tracking and accelerometers in the HMD. The user was able to interact with the mixed reality model through a combination of verbal, gestural, and haptic inputs. The proposed system was able to generate virtual female characters of different ethnicity to represent the patient. During the pilot experiments conducted in this study, the facial expression of the mixed reality human showed discomfort during any touching of the breast, and pain when the user pressed in an area that was designated as painful.

In 2012, Rivera-Gutierrez et al. [105] developed a physical virtual patient for medical students to conduct ophthalmic exams. The proposed system was based on a concept called Shader Lamps Virtual Patients (SLVP) where combination of shader lamps avatars and conversational virtual patients, as a means to achieve such physical presence were utilized. The physical system consisted of a styrofoam head with retro-reflective markers attached on the back, a pan-tilt unit for head motion control, and a static manikin body. A virtual character's head realized using virtual human software was projected onto the styrofoam head using a front-mounted Mitsubishi XD300U projector (1024 X 768 resolution). To correctly register the head of the virtual character onto the physical styrofoam head, a closed-loop process using the pan-tilt unit and an Eight-Natural-Point OptiTrack infrared camera tracking system was used. Results from the study suggested that proposed SLVP has an advantage over a SP because it can simulate pathological conditions such as restricted motion of one eye which a person with healthy eyes cannot exhibit.

The most recent hybrid approach was the physical virtual patient simulator by Daher et al. in 2020 [106]. An interchangeable translucent plastic shell was used as the physical structure of the patient body and face. And a virtual patient was rear-projected onto the shell using 2 AAXA P300 Pico projectors which provide imagery for the patient's head and body, respectively. The proposed physical-virtual patient simulator could display a range of multi sensory cues, including visual cues such as capillary refill, facial expressions, appearance changes as well as auditory and tactile cues.

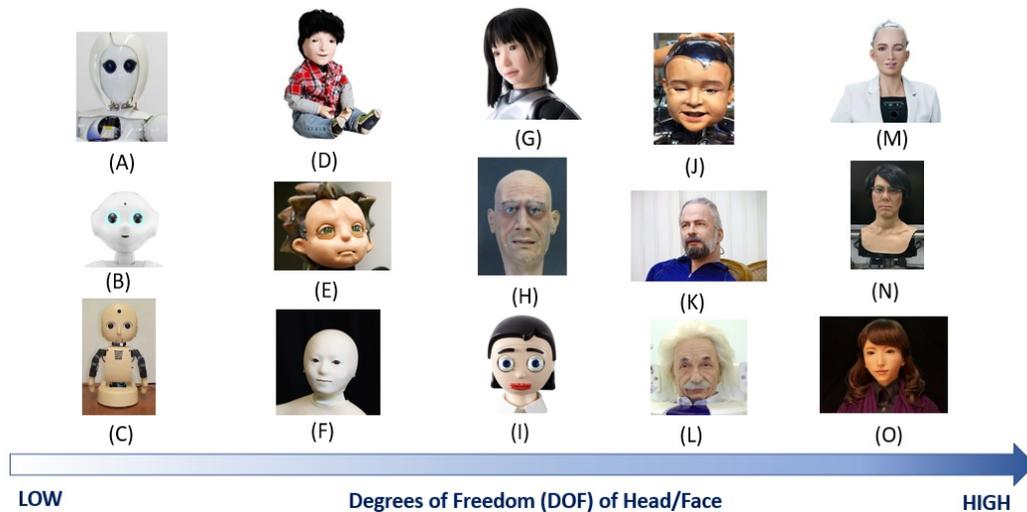
## IV. A BRIEF REVIEW ON FACIAL EXPRESSION RENDERING SYSTEMS REPORTED OUTSIDE MEDICAL TRAINING SIMULATOR DOMAIN

Research projects and insights from other fields such as robotics and computer graphics can also contribute to the developments of facial expression rendering systems in medical training simulators. This section reviews studies reported on virtual, physical and hybrid facial expression rendering approaches in those other research fields and discusses approaches and features that may be applicable to current and future developments of medical training simulators.

### A. PHYSICAL FACIAL EXPRESSION RENDERING APPROACHES

Extensive amount of work have been carried out in the field of robotics to develop robots with human-like faces or facial expressions. The majority of these work are related to social robots or human-robot interaction (HRI) domains. Fig. 6 shows physical robotic systems that can synthesise facial expressions from abstract, low fidelity to human-like, high fidelity levels.

The number of facial expressions these robots can simulate are directly proportional to the degrees of freedom (DOFs) they have. For example, AILA [108], Pepper Robot [109] and M3-Neony [110] have low numbers of DOFs and can only



**FIGURE 6.** Examples of physical facial expression rendering systems reported in Robotics. (A) AILA [108], (B) Pepper robot [109], (C) M3-Neony [110], (D) KASPAR [111], (E) Zeno [112], (F) Telenoid R1 [113], (G) HRP-4C [114], (H) RoMAN [115], (I) Flobi [116], (J) Diego-San [117], (K) Philip K Dick Research Robot [118], (L) HUBO [119], (M) Sophia [120], (N) Geminoid HI-5 [121], (O) Erica [122]

execute basic head and eye movements. Some researchers explored the possibility of synthesising facial expressions using cartoon-like robotic faces such as Zeno [112] and Flobi [116]. Though the appearance of these robots resemble fictional characters, their facial expressions and emotions were similar to a human's. This fictional character design approach allows features such as modularity to be added. For instance Flobi's modular structure allows easy access to the underlying hardware, and the visual features of the robot such as hairstyle and facial features can be altered easily.

Robots with high numbers of DOFs such as Sophia [120], Erica [122], HUBO [119], Geminoid HI-5 [121], PKD [118] and Diego-San [117] have greater expressiveness and more human-like appearances. In most designs [114], [117], [119], [120] servo motors were used as the actuators and in other robots [122], [121] actuation was realized using pneumatic actuators. The skins of the majority of the more human-like robots were made out of elastomer materials such as Frubber [120] and silicone rubber [121], [122], making their surface finishes visually similar to human skin.

Mechanical humanoid robots give great sense of realism but many may fall into the uncanny valley, a sudden decrease of the sense of affinity prompted by the realisation that the user is interacting with a robot [123]. The "creepiness" and uneasy feeling induced by this realisation may have drastic effect on the perception of facial expressions generated by the robot. Some robot designs tackle this problem by focusing on generating human-like facial expressions and emotions rather than replicating the human-like appearance. For instance the social robot Kismet [124] has an interactive emotion system and is capable of displaying multiple facial expression, but its appearance remains mostly mechanical.

## B. VIRTUAL FACIAL EXPRESSION RENDERING APPROACHES

Advancement in graphical processing and real-time rendering enabled researchers to build virtual human avatars with realistic and detailed facial features. To animate the facial expression of a virtual human avatar, both the morphological and dynamic characteristics of the virtual face must be considered [125]. Facial expressions generation of virtual avatars were based on methods either consist of exploiting the empirical and theoretical research such as FACS or the study of annotated corpus containing the expressions of emotions displayed by humans and virtual characters [125]. Many of the software solutions used FACS based methods and were made open-source to allow easy access for researchers and developers. Table 4 lists reported FACS-based software platforms for generating facial expressions.

**TABLE 4.** Open Source Applications and Platforms for Virtual Facial Expression Rendering

Software Platform	Number of controllable AUs
openFACS [126]	18
FACe! [127]	66
FACSGen 1.0 and 2.0 [128] [129]	58
HapFACS 1.0 and 3.0 [130] [131]	49
FACSHuman [132]	72

FACSe! [127] was the first 3-D face rendering software developed to use FACS to define bone system movements. It was built on the 3DSTUDIO MAX platform and controlled the bone rig of the face to change the activation intensity of different AUs to generate different facial expressions. Inspired by this application, FACSGen [128], [129] and HapFACS [130], [131] were developed with more functions including dynamic facial expressions that can be exported

TABLE 5. Examples of Hybrid Facial Expression Rendering Approaches

System/Study	System Snapshot	Virtual System Implementation	Physical System Implementation
Furhat [133]		Custom wide-angle (180 deg) projector (rear projection).	3 DOF head movements (Pan, tilt and roll).
Socibot [134]		Rear-fitted internal projector head with LED pico projector (rear projection).	3 DOF head movements (Pan, tilt and roll).
Mask-bot [135]		LED projector with a fish-eye conversion lens mounted behind the face mask (rear projection).	2 DOF head movements (Pan and tilt).
Bermano et al., 2013 [136]		3 Projectors and 5 cameras (front projection).	Deformable silicone skin attached to an underlying rigid articulated structure with 13 DOF and is driven by a set of motors.

as video clip, and bilateral AU activation. Game engines and computer graphics (CG) improved drastically in recent years, leading to the development of OpenFACS [126] and FACSHuman [132], which were built using gaming engines and CG software. OpenFACS was built with Unreal Engine [137] and could be used in multiple operating systems. Its human face rendering is very realistic, and offers the functions FACSGen and HapFACS have. However, a high performance computer is needed to meet the graphical processing requirement of the application or the frame rate drops significantly and the software becomes unresponsive. FACSHuman is a FACS-based plug-in for Makehuman [138], a Python-based human model generation software emerged from Blender [139]. It offers all the functions those other applications have and allows the human face models to be exported in many 3-D object formats.

There are many facial expression databases to help researchers to train, build and validate their models. Some databases only have static images, such as the Karolinska Directed Emotional Faces [140], which includes 4900 pictures of human facial expressions; and the Cohn-Kanade AU-Coded Facial Expression Database [141], which has 500 image sequences with FACS AU annotation and emotion-specific expressions. Other databases have images and videos. The MMI Facial Expression Database [142] contains 2900 videos and still images of 75 subjects, with AUs annotated on frame-level. The UNBC-McMaster Shoulder Pain Expression Archive Database [96] has 200 video sequences, 48398 FACS coded frames and associated pain scores with self-report and observer measures.

### C. HYBRID FACIAL EXPRESSION RENDERING APPROACHES

Lastly, Table 5 lists some of the hybrid facial expression systems that have both virtual and physical components.

Furhat [133] is a social robot that synthesises facial expression by having a virtual face projected onto a physical shell using a custom wide-angle projector. The physical head with the shell has 3 DOFs: pan, tilt and roll, which are controlled by servo motors. Using a 3-D animation software, Furhat is capable of synthesising faces of different ethnicity and gender. Similar hybrid approach was used in Socibot [134] and Mask-bot [135] which have 3 and 2 DOFs head movements in the physical system respectively. In all three aforementioned works, rear-projection method (a projector is mounted behind a physical shell) is used with no active components in the physical shell.

Bermano et al. [136] proposed an augmented physical avatar using projector-based illumination. This model has a high number of DOFs, and as mentioned previously the number of expressions physical systems can produce are directly proportional but also limited by the DOFs they have. To make this model synthesise more expressions, the authors introduced a system that decomposed the motion into low-frequency motions that were physically reproduced by the physical robotic head, and high-frequency details that were added using projected shading. The physical head was implemented using a layer of silicone skin attached to an underlying rigid articulated structure with 13 DOFs actuated by a set of motors. 5 cameras and 3 projectors were mounted in front of the physical face to reconstruct and project the virtual face. The authors in this study claimed that result of their system was a highly expressive physical avatar that features facial details and motion otherwise unattainable due to physical constraints.

### V. DISCUSSION AND FUTURE DIRECTIONS

Given a haptic task such as physical examination, our previous work showed that the brain takes two broad classes of actions – one to condition the body such as searching

for the best finger stiffness [143], and the other to control movements and forces [144], [145]. In the case of physical examination of patients, the process of combining visual perception with haptic feedback plays a critical role in the motor control pathways. For instance, subjective interpretation of facial reactions during physical examination such as palpation can lead to reduced levels of indentation. This may require a different finger speed regulation method to gain haptic information about a physiological condition. Misperception of constraints associated with palpation behavioural variables can therefore lead to sub-optimal diagnostic methods. Medical training poses the challenge of finding best design methods for patient simulators to prompt the learning of multi-modal sensor integration and sensory-motor coordination.

The face is one of the most expressive areas of the body which is capable of producing about twenty thousand different facial expressions [146]. Naturally, this makes facial expressions of patients one of the powerful non-verbal feedback modalities that medical personnel should account for during medical training and practice. As such, this paper reviewed the developments and current implementations of facial expression rendering approaches in medical training simulators. We reviewed models describing facial expression generation, explained the advantages and significance in analysing facial expressions in clinical practice, highlighted the relationship between pain reaction and facial expressions, and introduced systems and frameworks for classifying facial expressions.

### A. DESIGN CHALLENGES

We highlighted the debates around the psychological basis of tracing facial expressions to pain. It became clear that the context cannot be separated from the facial expression to interpret them. In medical education, the patients' medical history and cultural background are important factors to interpret, in conjunction with facial expressions to diagnose a physiological condition. Moreover, other feedback modalities such as vocal and muscle tension also provide useful clues together with facial expressions. Therefore, potential simulators should be able to present a fair variety of medical contexts in order to give a well rounded preliminary training to trainees for generalizable skills.

We then reviewed medical training simulators with facial expression rendering systems and classified them to three categories: physical, virtual and hybrid. Their application domains, commercial availability, implementation methods, and facial expression simulation abilities were evaluated. Facial expression rendering approaches and systems in other research fields were also discussed, as some of the proposed methods and insights can be transferred to develop such systems for medical training simulators. It became clear to us that hybrid approaches have the advantage of being able to achieve a higher level of agency while offering the flexibility to present a variety of patient contexts. This is very important because empathetic connection is an important

factor during medical examinations. Ethnic appearance is one such important contextual factor that can be easily changed in a hybrid approach. At a higher level of complexity, it would be interesting to explore how the physical morphology of a hybrid patient simulator can be controlled to render gender and ethnicity interactions.

The challenge for designers would be to maximise the efficacy of such designs using material, actuators, and graphics rendering tools to manage the level of agency of the medical training simulator with facial expression rendering without crossing over to the uncanny valley. Certain medical scenarios especially those involving the examination of the face itself can pose extra challenges for designers. For example in a dental or emergency treatment training simulator, the mouth of the simulated patient plays an important role and the face is directly involved in the training process. Fig. 7 gives a list of such considerations for all stake holders: medical educators, students and simulator developers.

To further improve the realism of the training experience, the robotic system could utilise multimodal fusion by capturing the state of the examiner, such as facial expressions, speech, force profiles and body movements. Sentic blending strategies [147] can be used to integrate quantitative data such as the force profile with abstract information such as facial expressions and speech to derive how the examiner feels at any given time. As mentioned previously, intrinsic factors affect how the patient reacts to pain, and the cognitive and affective information received by the patient is affected by the state of the examiner. Cameras and microphones can be used in the hybrid systems to capture the examiner's movements, facial expressions and speech data, and these information can be projected into a common AffectiveSpace through a fuzzy classifier [148]. Future development of hybrid systems may benefit from sentic blending and multimodal fusion to improve the realism of the training experience, and to capture more insightful information of the student's performance.

### B. FUTURE DIRECTIONS

This review highlights the need to have controllable medical training simulators such as robotic patient simulators that can present multiple physiological conditions of patients from different gender and ethnic backgrounds. This will allow instructors to assist medical students to learn robust techniques and efficient methods to combine visual and haptic feedback when performing medical procedures on patients from diverse backgrounds. For instance, in the case of a physical examination simulator, this requires such robotic simulators to consist of 1) controllable sensorized internal organs that can present several classes of symptoms and measure palpation behaviours, 2) a facial expression rendering method to present essential features of visual feedback relevant to the training context, and 3) methods to quantify trainee's examination behaviors to provide focused feedback for peer assisted learning [149] and to optimize a given set of training criteria.

### Considerations and needs for:

#### Medical Educators

- Demonstrations and Teaching
- Progress monitoring of the students
- Examination and evaluations
- Training difficulty control
- Calibration

#### Simulator Developers

- Cost of implementation
- Complexity of implementation
- Safety aspect of the simulator
- Application area of the simulator
- Level of the realism
- Flexibility of changing demographic features

#### Medical Students

- Ease of use
- Physicality
- Group learning
- Accurate and real-time feedback
- Accessibility and independent learning

FIGURE 7. Considerations and needs for three stakeholders of medical training simulators.

## VI. CONCLUSION

Being able to present a variety of patient contexts is an open design challenge for patient simulators. Out of the currently available methods, physical simulators provide accurate haptic feedback but lacks patient customisation and detailed facial expression simulation. Virtual systems on the other hand provide facial expression simulations with fine detail and realism, and can simulate patients with different demographic attributes, but they often do not provide haptic feedback or interface with hardware sensors to take physical inputs. Hybrid systems integrate features from both physical and virtual systems and are capable of delivering accurate haptic and visual feedback. Based on this review we conclude that there are several opportunities for technology advances to maximise the efficacy of patient simulators with facial expressions. New methods can be found to control the physical morphology of the robot to render gender and ethnic diversity. This can be backed by animated solutions to overlay detailed nuances representing the patients' medical and cultural context. This poses an opportunity for AI techniques such as deep neural networks to parameterize complex linguistic phenomena such as culture, gender, and ethnic interactions. In this regard, considerations to determine the responsibility sharing between the physical and virtual rendering methods has become a key future challenge.

## REFERENCES

- [1] C. Vincent, G. Neale, and M. Woloshynowych, "Adverse events in british hospitals: Preliminary retrospective record review," *BMJ (Clinical research ed.)*, vol. 322, pp. 517–9, 04 2001.
- [2] L. Kohn, J. Corrigan, and M. Donaldson, *To Err is Human: Building a Safer Health System*. National academy press Washington, DC, 01 2000, vol. 6.
- [3] M. F. Charles, *Inspection, Auscultation, Palpation, and Percussion of the Abdomen*, 1990, vol. 3.
- [4] Y. Zhang, R. Phillips, J. Ward, and S. Pisharody, *A survey of simulators for palpation training*, Jan. 2009, vol. 142, pp. 444 – 446.
- [5] H. S. Barrows, "Simulated patients in medical teaching," *Canadian Medical Association Journal*, vol. 98, no. 14, p. 674, 1968.
- [6] D. A. Davis, M. A. Thomson, A. D. Oxman, and R. B. Haynes, "Changing physician performance: a systematic review of the effect of continuing medical education strategies," *JAMA*, vol. 274, no. 9, pp. 700–705, 1995.
- [7] R. Likic and S. R. Maxwell, "Prevention of medication errors: teaching and training," *British journal of clinical pharmacology*, vol. 67, no. 6, pp. 656–661, 2009.
- [8] S. Maloney, M. Storr, S. Paynter, P. Morgan, and D. Ilic, "Investigating the efficacy of practical skill teaching: a pilot-study comparing three educational methods," *Advances in Health Sciences Education*, vol. 18, no. 1, pp. 71–80, 2013.
- [9] K. M. Prkachin, S. Berzins, and S. R. Mercer, "Encoding and decoding of pain expressions: a judgement study," *Pain*, vol. 58, no. 2, pp. 253–259, 1994.
- [10] J. D. Greenspan, R. M. Craft, L. LeResche, L. Arendt-Nielsen, K. J. Berkley, R. B. Fillingim, M. S. Gold, A. Holdcroft, S. Lautenbacher, E. A. Mayer et al., "Studying sex and gender differences in pain and analgesia: a consensus report," *Pain*, vol. 132, pp. S26–S45, 2007.
- [11] L. Kacperck, "Non-verbal communication: the importance of listening," *British Journal of Nursing*, vol. 6, no. 5, pp. 275–279, 1997.
- [12] A. Mehrabian et al., *Silent messages*. Wadsworth Belmont, CA, 1971, vol. 8, no. 152.
- [13] A. Pease, *Body language: How to read others' thoughts by their gestures*. Sheldon Press, 1984.
- [14] D. Rosenstein and H. Oster, "Differential facial responses to four basic tastes in newborns," *Child development*, pp. 1555–1568, 1988.
- [15] P. Ekman, "Facial expression and emotion," *American psychologist*, vol. 48, no. 4, p. 384, 1993.
- [16] K. M. Prkachin and K. D. Craig, "Expressing pain: the communication and interpretation of facial pain signals," *Journal of Nonverbal Behavior*, vol. 19, no. 4, pp. 191–205, 1995.
- [17] K. M. Hoffman, S. Trawalter, J. R. Axt, and M. N. Oliver, "Racial bias in pain assessment and treatment recommendations, and false beliefs about biological differences between blacks and whites," *Proceedings of the National Academy of Sciences*, vol. 113, no. 16, pp. 4296–4301, 2016.
- [18] E. A. Jacobs, I. Rolle, C. E. Ferrans, E. E. Whitaker, and R. B. Warnecke, "Understanding african americans' views of the trustworthiness of physicians," *Journal of general internal medicine*, vol. 21, no. 6, p. 642, 2006.
- [19] J. M. Weller, D. Nestel, S. D. Marshall, P. M. Brooks, and J. J. Conn, "Simulation in clinical teaching and learning," *Medical Journal of Australia*, vol. 196, no. 9, pp. 594–594, 2012.
- [20] A. Ziv, P. R. Wolpe, S. D. Small, and S. Glick, "Simulation-based medical education: an ethical imperative," *Simulation in Healthcare*, vol. 1, no. 4, pp. 252–256, 2006.
- [21] L. He, N. Herzig, S. de Lusignan, and T. Nanayakkara, "Granular jamming based controllable organ design for abdominal palpation," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2018, pp. 2154–2157.
- [22] A. Ziv, S. Ben-David, and M. Ziv, "Simulation based medical education: an opportunity to learn from errors," *Medical teacher*, vol. 27, no. 3, pp. 193–199, 2005.
- [23] S. McLaughlin, S. Clarke, S. Menon, T. Noeller, Y. Okuda, M. Smith, and C. Strother, *The Comprehensive Textbook of Healthcare Simulation*, 04 2013.
- [24] H. Barrows, "An overview of the uses of standardized patients for teaching and evaluating clinical skills. aamc," *Academic medicine : journal of the Association of American Medical Colleges*, vol. 68, no. 6, p. 443–51; discussion 451–3, June 1993.
- [25] T. C. Tsai, "Using children as standardised patients for assessing clinical competence in paediatrics," *Archives of Disease in Childhood*, vol. 89, no. 12, pp. 1117–1120, 2004.
- [26] B. Jackson and L. Rosenberg, "Force feedback and medical simulation," *Interactive Technology and the New Paradigm for Healthcare*, Morgan et al. (Eds), pp. 147–151, 1995.
- [27] "Multi-Venous IV Training Arms," accessed 01-06-2020. [Online]. Available: <https://www.laerdal.com/gb/products/skills-proficiency/venous-arterial-access/multi-venous-iv-training-arms/>

- [28] "SimMan 3G," accessed 17-05-2020. [Online]. Available: <https://www.laerdal.com/gb/products/simulation-training/emergency-care-trauma/simman-3g/>
- [29] C. M. Pugh, K. B. Iannitelli, D. Rooney, and L. Salud, "Use of mannequin-based simulation to decrease student anxiety prior to interacting with male teaching associates," *Teaching and learning in medicine*, vol. 24, no. 2, pp. 122–127, 2012.
- [30] W. C. McGaghie, S. B. Issenberg, M. E. R. Cohen, J. H. Barsuk, and D. B. Wayne, "Does simulation-based medical education with deliberate practice yield better results than traditional clinical education? a meta-analytic comparative review of the evidence," *Academic medicine: journal of the Association of American Medical Colleges*, vol. 86, no. 6, p. 706, 2011.
- [31] T. R. Coles, D. Meglan, and N. W. John, "The role of haptics in medical training simulators: A survey of the state of the art," *IEEE Transactions on Haptics*, vol. 4, no. 1, pp. 51–66, 2011.
- [32] D. Escobar-Castillejos, J. Noguez, L. Neri, A. Magana, and B. Benes, "A review of simulators with haptic devices for medical training," *Journal of medical systems*, vol. 40, no. 4, p. 1–22, Apr. 2016.
- [33] M. Kunz, S. Scharmann, U. Hemmeter, K. Schepelmann, and S. Lautenbacher, "The facial expression of pain in patients with dementia," *PAIN®*, vol. 133, no. 1-3, pp. 221–228, 2007.
- [34] J. Russell and L. Barrett, "Core affect, prototypical emotional episodes, and other things called emotion: Dissecting the elephant," *Journal of personality and social psychology*, vol. 76, pp. 805–19, 06 1999.
- [35] K. Grabowski, A. Rynkiewicz, A. Lassalle, S. Baron-Cohen, B. Schuller, N. Cummins, A. Baird, J. Podgórska-Bednarz, A. Pieniążek, and I. Lucka, "Emotional expression in psychiatric conditions: New technology for clinicians," *Psychiatry and clinical neurosciences*, vol. 73, no. 2, pp. 50–62, 2019.
- [36] J. Kiverstein and M. Miller, "The embodied brain: Towards a radical embodied cognitive neuroscience," *Frontiers in Human Neuroscience*, vol. 9, 05 2015.
- [37] L. F. Barrett, R. Adolphs, S. Marsella, A. M. Martinez, and S. D. Pollak, "Emotional expressions reconsidered: Challenges to inferring emotion from human facial movements," *Psychological Science in the Public Interest*, vol. 20, no. 1, pp. 1–68, 2019.
- [38] L. Barrett, "Are emotions natural kinds?" *Perspectives on Psychological Science*, vol. 1, pp. 28–58, 03 2006.
- [39] D. T. Cordaro, R. Sun, D. Keltner, S. Kamble, N. Huddar, and G. McNeil, "Universals and cultural variations in 22 emotional expressions across five cultures," *Emotion (Washington, D.C.)*, vol. 18, no. 1, p. 75–93, February 2018.
- [40] D. Matsumoto, D. Keltner, M. N. Shiota, M. O'Sullivan, and M. Frank, *Facial expressions of emotion.. Handbook of emotions*, 3rd ed., 2008.
- [41] L. LeResche, "Facial expression in pain: a study of candid photographs," *Journal of Nonverbal Behavior*, vol. 7, no. 1, pp. 46–56, 1982.
- [42] P. Ekman and W. V. Friesen, "Facial action coding system: a technique for the measurement of facial movement," 1978.
- [43] K. A. Lindquist, T. D. Wager, H. Kober, E. Bliss-Moreau, and L. F. Barrett, "The brain basis of emotion: a meta-analytic review," *The Behavioral and brain sciences*, vol. 35, no. 3, p. 121–143, June 2012.
- [44] K. Vytal and S. Hamann, "Neuroimaging support for discrete neural correlates of basic emotions: a voxel-based meta-analysis," *Journal of cognitive neuroscience*, vol. 22, no. 12, p. 2864–2885, December 2010.
- [45] R. Calvo and S. D'Mello, "Affect detection: An interdisciplinary review of models, methods, and their applications," *T. Affective Computing*, vol. 1, pp. 18–37, 01 2010.
- [46] L. F. Barrett, *How emotions are made: The secret life of the brain*, 2017.
- [47] K. Lindquist and L. Barrett, "A functional architecture of the human brain: Emerging insights from the science of emotion," *Trends in cognitive sciences*, vol. 16, pp. 533–40, 10 2012.
- [48] L. Barrett and E. Bliss-Moreau, "Affect as a psychological primitive," *Advances in experimental social psychology*, vol. 41, pp. 167–218, 12 2009.
- [49] L. Barrett, "The theory of constructed emotion: an active inference account of interoception and categorization," *Social Cognitive and Affective Neuroscience*, vol. 12, 05 2017.
- [50] M. Miller and A. Clark, "Happily entangled: prediction, emotion, and the embodied mind," *Synthese*, vol. 195, 04 2017.
- [51] J. K. Bone, G. Lewis, K. S. Button, L. Duffy, C. J. Harmer, M. R. Munafò, I. S. Penton-Voak, N. J. Wiles, and G. Lewis, "Variation in recognition of happy and sad facial expressions and self-reported depressive symptom severity: A prospective cohort study," *Journal of affective disorders*, vol. 257, pp. 461–469, 2019.
- [52] D. E. Schoth and C. Liossi, "Biased interpretation of ambiguous information in patients with chronic pain: A systematic review and meta-analysis of current studies," *Health Psychology*, vol. 35, no. 9, p. 944, 2016.
- [53] P. Ekman and W. V. Friesen, "Constants across cultures in the face and emotion," *Journal of personality and social psychology*, vol. 17, no. 2, p. 124, 1971.
- [54] C. Izard, C. Malatesta, and J. Osofsky, "Handbook of infant development," 1987.
- [55] S. Tomkins, "The positive affects. affect, imagery, consciousness, vol. 1," 1962.
- [56] C. Chen, C. Crivelli, O. G. Garrod, P. G. Schyns, J.-M. Fernández-Dols, and R. E. Jack, "Distinct facial expressions represent pain and pleasure across cultures," *Proceedings of the National Academy of Sciences*, vol. 115, no. 43, pp. E10 013–E10 021, 2018.
- [57] U. S. Donges, A. Kersting, and T. Suslow, "Women's greater ability to perceive happy facial emotion automatically: gender differences in affective priming," *PloS one*, vol. 7, no. 7, 2012.
- [58] D. M. Isaacowitz, C. E. Löckenhoff, R. D. Lane, R. Wright, L. Sechrest, R. Riedel, and P. T. Costa, "Age differences in recognition of emotion in lexical stimuli and facial expressions," *Psychology and aging*, vol. 22, no. 1, p. 147, 2007.
- [59] J. E. Kilbride and M. Yarczower, "Ethnic bias in the recognition of facial expressions," *Journal of Nonverbal Behavior*, vol. 8, no. 1, pp. 27–41, 1983.
- [60] R. A. Sternbach, "Psychophysiology of pain," *The International Journal of Psychiatry in Medicine*, vol. 6, no. 1-2, pp. 63–73, 1975.
- [61] C. M. Ferguson, "Inspection, auscultation, palpation, and percussion of the abdomen," *Clinical Methods: The History, Physical and Laboratory Examinations*. 3rd edn. Boston, MA: Butterworths, 1990.
- [62] K. Herr, P. J. Coyne, M. McCaffery, R. Manworren, and S. Merkel, "Pain assessment in the patient unable to self-report: position statement with clinical practice recommendations," *Pain Management Nursing*, vol. 12, no. 4, pp. 230–250, 2011.
- [63] E. K. Breivik, G. A. Björmsen, and E. Skovlund, "A comparison of pain rating scales by sampling from clinical trial data," *The Clinical journal of pain*, vol. 16, no. 1, pp. 22–28, 2000.
- [64] C. L. Hicks, C. L. von Baeyer, P. A. Spafford, I. van Korlaar, and B. Goodenough, "The faces pain scale-revised: toward a common metric in pediatric pain measurement," *Pain*, vol. 93, no. 2, pp. 173–183, 2001.
- [65] A. K. Agrawal and J. Feusner, "Chapter 33 - supportive care of patients with cancer," in *Lanzkowsky's Manual of Pediatric Hematology and Oncology (Sixth Edition)*, sixth edition ed., P. Lanzkowsky, J. M. Lipton, and J. D. Fish, Eds. San Diego: Academic Press, 2016, pp. 620 – 655.
- [66] K. R. Scherer, *Handbook of methods in nonverbal behavior research*. Cambridge University Press, 1985.
- [67] P. Ekman, "Expression and the nature of emotion," *Approaches to emotion*, vol. 3, no. 19, p. 344, 1984.
- [68] W. P. Chapman and C. M. Jones, "Variations in cutaneous and visceral pain sensitivity in normal subjects," *The Journal of clinical investigation*, vol. 23, no. 1, pp. 81–91, 1944.
- [69] M. Sown, "A preliminary note on pattern recognition of facial emotional expression," in *The 4th International Joint Conferences on Pattern Recognition, 1978, 1978*.
- [70] D. Terzopoulos and K. Waters, "Analysis of facial images using physical and anatomical models," in *Proceedings Third International Conference on Computer Vision*, 1990, pp. 727–732.
- [71] K. Mase, "Recognition of facial expression from optical flow," *IEICE TRANSACTIONS on Information and Systems*, vol. 74, no. 10, pp. 3474–3483, 1991.
- [72] J. G. Neely, J. Y. Cheung, M. Wood, J. Byers, and A. Rogerson, "Computerized quantitative dynamic analysis of facial motion in the paralyzed and synkinetic face," *The American journal of otology*, vol. 13, no. 2, pp. 97–107, 1992.
- [73] D. A. Wood, G. B. Hughes, M. Secic, and T. L. Good, "Objective measurement of normal facial movement with video microscaling," *The American journal of otology*, vol. 15, no. 1, pp. 61–65, 1994.
- [74] Y. Yacoob and L. S. Davis, "Recognizing human facial expressions from long image sequences using optical flow," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 18, no. 6, pp. 636–642, 1996.
- [75] C. Padgett, G. Cottrell, and R. Adolphs, "Categorical perception in facial emotion classification," in *Proceedings of the Eighteenth Annual Conference of the Cognitive Science Society*, vol. 18, 1996, p. 249.
- [76] M. Rosenblum, Y. Yacoob, and L. S. Davis, "Human expression recognition from motion using a radial basis function network architecture,"

- IEEE transactions on neural networks*, vol. 7, no. 5, pp. 1121–1138, 1996.
- [77] K. Yuen, I. Inokuchi, M. Maeta, S.-I. Kawakami, and Y. Masuda, “Evaluation of facial palsy by moiré topography index,” *Otolaryngology—Head and Neck Surgery*, vol. 117, no. 5, pp. 567–572, 1997.
- [78] I. A. Essa and A. P. Pentland, “Coding, analysis, interpretation, and recognition of facial expressions,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 19, no. 7, pp. 757–763, 1997.
- [79] K. Fukui and O. Yamaguchi, “Facial feature point extraction method based on combination of shape extraction and pattern matching,” *Systems and Computers in Japan*, vol. 29, no. 6, pp. 49–58, 1998.
- [80] J. F. Cohn, A. J. Zlochower, J. Lien, and T. Kanade, “Automated face analysis by feature point tracking has high concurrent validity with manual face coding,” *Psychophysiology*, vol. 36, no. 1, pp. 35–43, 1999.
- [81] G. Donato, M. S. Bartlett, J. C. Hager, P. Ekman, and T. J. Sejnowski, “Classifying facial actions,” *IEEE Transactions on pattern analysis and machine intelligence*, vol. 21, no. 10, pp. 974–989, 1999.
- [82] Z. Zhang, “Feature-based facial expression recognition: Sensitivity analysis and experiments with a multilayer perceptron,” *International journal of pattern recognition and Artificial Intelligence*, vol. 13, no. 06, pp. 893–911, 1999.
- [83] J. J.-J. Lien, T. Kanade, J. F. Cohn, and C.-C. Li, “Detection, tracking, and classification of action units in facial expression,” *Robotics and Autonomous Systems*, vol. 31, no. 3, pp. 131–146, 2000.
- [84] C. Izard, “The maximally discriminative facial movement coding system (max),” 1979.
- [85] C. Izard, L. Dougherty, and E. Hembree, “A system for identifying affect expressions by holistic judgements,” 1983.
- [86] P. Dulguerov, F. Marchal, D. Wang, and C. Gysin, “Review of objective topographic facial nerve evaluation methods,” *American Journal of Otolaryngology*, vol. 20, no. 5, pp. 672–8, 1999.
- [87] J. W. House and D. E. Brackmann, “Facial nerve grading system,” *Journal of American Academy of Otolaryngology-Head and Neck Surgery*, vol. 93, pp. 184–193, 1985.
- [88] K. M. Prkachin, “The consistency of facial expressions of pain: a comparison across modalities,” *Pain*, vol. 51, no. 3, pp. 297–306, 1992.
- [89] K. D. Craig and C. J. Patrick, “Facial expression during induced pain,” *Journal of personality and social psychology*, vol. 48, no. 4, p. 1080, 1985.
- [90] K. D. Craig, S. A. Hyde, and C. J. Patrick, “Genuine, suppressed and faked facial behavior during exacerbation of chronic low back pain,” *Pain*, vol. 46, no. 2, pp. 161–171, 1991.
- [91] C. J. Patrick, K. D. Craig, and K. M. Prkachin, “Observer judgments of acute pain: Facial action determinants,” *Journal of personality and social psychology*, vol. 50, no. 6, p. 1291, 1986.
- [92] L. LeResche and S. F. Dworkin, “Facial expressions of pain and emotions in chronic tmd patients,” *Pain*, vol. 35, no. 1, pp. 71–78, 1988.
- [93] K. M. Prkachin and S. R. Mercer, “Pain expression in patients with shoulder pathology: validity, properties and relationship to sickness impact,” *Pain*, vol. 39, no. 3, pp. 257–265, 1989.
- [94] D. Simon, K. D. Craig, F. Gosselin, P. Belin, and P. Rainville, “Recognition and discrimination of prototypical dynamic expressions of pain and emotions,” *PAIN®*, vol. 135, no. 1-2, pp. 55–64, 2008.
- [95] K. M. Prkachin and P. E. Solomon, “The structure, reliability and validity of pain expression: Evidence from patients with shoulder pain,” *Pain*, vol. 139, no. 2, pp. 267–274, 2008.
- [96] P. Lucey, J. F. Cohn, K. M. Prkachin, P. E. Solomon, and I. Matthews, “Painful data: The unbc-mcmaster shoulder pain expression archive database,” in *Face and Gesture 2011*. IEEE, 2011, pp. 57–64.
- [97] “Pediatric HAL,” accessed 17-05-2020. [Online]. Available: <https://www.gaumard.com/s2225>
- [98] “SIMROID,” accessed 17-05-2020. [Online]. Available: <https://www.morita.com/group/en/products/educational-and-training-systems/training-simulation-system/simroid/>
- [99] “DENTAROID,” accessed 17-05-2020. [Online]. Available: <https://www.nissin-dental.net/products/DentalTrainingProducts/ Dental-Simulator/dentaroid/index.html>
- [100] E. Baldrighi, N. Thayer, M. Stevens, S. R. Echols, and S. Priya, “Design and implementation of the bio-inspired facial expressions for medical mannequin,” *International Journal of Social Robotics*, vol. 6, no. 4, pp. 555–574, 2014.
- [101] L. D. Wandner, L. A. Stutts, A. F. Alqudah, J. G. Craggs, C. D. Scipio, A. T. Hirsh, and M. E. Robinson, “Virtual human technology: patient demographics and healthcare training factors in pain observation and treatment recommendations,” *Journal of pain research*, vol. 3, p. 241, 2010.
- [102] M. Moosaei, M. J. Gonzales, and L. D. Riek, “Naturalistic pain synthesis for virtual patients,” in *International Conference on Intelligent Virtual Agents*. Springer, 2014, pp. 295–309.
- [103] K. Maicher, D. Danforth, A. Price, L. Zimmerman, B. Wilcox, B. Liston, H. Cronau, L. Belknap, C. Ledford, D. Way, D. Post, A. Macerollo, and M. Rizer, “Developing a conversational virtual standardized patient to enable students to practice history-taking skills,” *Simulation in Healthcare*, vol. 12, no. 2, pp. 124–131, 2017.
- [104] A. Kotranza and B. Lok, “Virtual human + tangible interface = mixed reality human an initial exploration with a virtual breast exam patient,” in *2008 IEEE Virtual Reality Conference*, 2008, pp. 99–106.
- [105] R.-G. Diego, W. Greg, L. Peter, W. Mary, C. Juan, C. David A., F. Henry, and L. Benjamin, “Shader lamps virtual patients: the physical manifestation of virtual patients,” *Studies in health technology and informatics*, vol. 173, pp. 372–378, 2012.
- [106] S. Daher, J. Hochreiter, R. Schubert, L. Gonzalez, J. Cendan, M. Anderson, D. A. Diaz, and G. F. Welch, “The physical-virtual patient simulator,” *Simulation in Healthcare: The Journal of the Society for Simulation in Healthcare*, vol. 15, no. 2, pp. 115–121, 2020.
- [107] J. Buchanan, “Use of simulation technology in dental education,” *Journal of dental education*, vol. 65, no. 11, pp. 1225–1231, 2001.
- [108] J. Lemburg, J. de Gea Fernández, M. Eich, D. Mronga, P. Kampmann, A. Vogt, A. Aggarwal, Yuping Shi, and F. Kirchner, “Aila - design of an autonomous mobile dual-arm robot,” in *2011 IEEE International Conference on Robotics and Automation*, 2011, pp. 5147–5153.
- [109] “Pepper robot,” accessed 17-05-2020. [Online]. Available: <https://www.softbankrobotics.com/emea/en/pepper>
- [110] T. Minato, F. DallaLibera, S. Yokokawa, Y. Nakamura, H. Ishiguro, and E. Menegatti, “A baby robot platform for cognitive developmental robotics,” 2009.
- [111] K. Dautenhahn, C. L. Nehaniv, M. L. Walters, B. Robins, H. Kose-Bagci, N. A. Mirza, and M. Blow, “Kaspar – a minimally expressive humanoid robot for human–robot interaction research,” *Applied Bionics and Biomechanics*, vol. 6, no. 3-4, pp. 369–397, 2009.
- [112] D. Hanson, S. Baurmann, T. Riccio, R. Margolin, T. Dockins, M. Tavares, and K. Carpenter, “Zeno: A cognitive character,” in *AAAI Conference on Artificial Intelligence*, 2008.
- [113] K. Ogawa, S. Nishio, K. Koda, <br>Giuseppe Balistreri, T. Watanabe, and <br>and Hiroshi Ishiguro, “Exploring the natural reaction of young and aged person with telenoid in a real world,” *Journal of Advanced Computational Intelligence and Intelligent Informatics*, vol. 15, no. 5, pp. 592–597, 2011.
- [114] S. Kajita, K. Kenji, K. Fumio, H. Kensuke, M. Mitsuharu, N. Shin’ichiro, M. Kanako, F. Kiyoshi, N. Ee Sian, H. Isao, Y. Kazuhito, and H. Hirohisa, “Cybernetic human hrp-4c: A humanoid robot with human-like proportions,” in *Robotics Research. Springer Tracts in Advanced Robotics*, vol. 70, 2011, pp. 301–314.
- [115] J. Hirth and K. Berns, “Motives as intrinsic activation for human-robot interaction,” in *2008 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2008, pp. 773–778.
- [116] F. Hegel, F. Eyssel, and B. Wrede, “The social robot ‘flobi’: Key concepts of industrial design,” in *19th International Symposium in Robot and Human Interactive Communication*, 2010, pp. 107–112.
- [117] “Diego San,” accessed 17-05-2020. [Online]. Available: <https://www.hansonrobotics.com/diego-san/>
- [118] D. Hanson, A. Olney, S. Prillman, E. Mathews, M. Zielke, D. Hammons, R. Fernandez, and H. Stephanou, “Upending the uncanny valley,” in *Proceedings of the 20th National Conference on Artificial Intelligence - Volume 4*, 2005, p. 1728–1729.
- [119] J. Oh, D. Hanson, W. Kim, Y. Han, J. Kim, and I. Park, “Design of android type humanoid robot albert hubo,” in *2006 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006, pp. 1428–1433.
- [120] “Sophia,” accessed 17-05-2020. [Online]. Available: <https://www.hansonrobotics.com/sophia/>
- [121] “Geminoid HI-5,” accessed 17-05-2020. [Online]. Available: <http://www.geminoid.jp/en/robots.html>
- [122] D. F. Glas, T. Minato, C. T. Ishi, T. Kawahara, and H. Ishiguro, “Erica: The erato intelligent conversational android,” in *2016 25th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)*, 2016, pp. 22–29.

- [123] M. Mori, K. F. MacDorman, and N. Kageki, "The uncanny valley [from the field]," *IEEE Robotics & Automation Magazine*, vol. 19, no. 2, pp. 98–100, 2012.
- [124] C. Breazeal and J. Velásquez, "Toward teaching a robot 'infant' using emotive communication acts," in *Proceedings of the 1998 Simulated Adaptive Behavior Workshop on Socially Situated Intelligence*. Citeseer, 1998, pp. 25–40.
- [125] M. Ochs, R. Niewiadomski, and C. Pelachaud, "18 facial expressions of emotions for virtual characters," in *The Oxford Handbook of Affective Computing*. Oxford University Press, USA, 2015, p. 261.
- [126] V. Cuculo and A. D'Amelio, "Openfacs: An open source facs-based 3d face animation system," in *Image and Graphics*, Y. Zhao, N. Barnes, B. Chen, R. Westermann, X. Kong, and C. Lin, Eds. Cham: Springer International Publishing, 2019, pp. 232–242.
- [127] S. Villagrana and A. Susín Sánchez, "Face! 3d facial animation system based on facs," in *IV Iberoamerican symposium in computer graphics*, 2009, pp. 203–209.
- [128] E. B. Roesch, L. Tamarit, L. Reveret, D. Grandjean, D. Sander, and K. R. Scherer, "Facs-gen: A tool to synthesize emotional facial expressions through systematic manipulation of facial action units," *Journal of Nonverbal Behavior*, vol. 35, no. 1, pp. 1–16, 2011.
- [129] E. G. Krumhuber, L. Tamarit, E. B. Roesch, and K. R. Scherer, "Facs-gen 2.0 animation software: Generating three-dimensional facs-valid facial expressions for emotion research." *Emotion*, vol. 12, no. 2, p. 351, 2012.
- [130] R. Amini, U. Yasavur, and C. Lisetti, "Hapfacs 1.0: Software/api for generating facs-based facial expressions," in *Proceedings of the 3rd Symposium on Facial Analysis and Animation*, 2012.
- [131] R. Amini, C. Lisetti, and G. Ruiz, "Hapfacs 3.0: Facs-based facial expression generator for 3d speaking virtual characters," *IEEE Transactions on Affective Computing*, vol. 6, no. 4, pp. 348–360, 2015.
- [132] M. Gilbert, S. Demarchi, and I. Urdapilleta, "Facshuman a software to create experimental material by modeling 3d facial expression," in *Proceedings of the 18th International Conference on Intelligent Virtual Agents*, 2018, pp. 333–334.
- [133] "Furhat Robotics," accessed 17-05-2020. [Online]. Available: <https://furhatrobotics.com/>
- [134] "socibot," accessed 17-05-2020. [Online]. Available: <https://www.engineeredarts.co.uk/socibot/>
- [135] T. Kuratate, Y. Matsusaka, B. Pierce, and G. Cheng, "'mask-bot': A life-size robot head using talking head animation for human-robot communication," in *IEEE International Conference on Humanoid Robots*, 2011, pp. 99–104.
- [136] A. Bermanno, P. Brüscheweiler, A. Grundhöfer, D. Iwai, B. Bickel, and M. Gross, "Augmenting physical avatars using projector-based illumination," *ACM Trans. Graph.*, vol. 32, no. 6, Nov. 2013.
- [137] "Unreal Engine," accessed 17-05-2020. [Online]. Available: <https://www.unrealengine.com/en-US/>
- [138] "makehumancommunity." [Online]. Available: <http://www.makehumancommunity.org/>
- [139] "Blender," accessed 17-05-2020. [Online]. Available: <https://www.blender.org/>
- [140] D. Lundqvist, A. Flykt, and A. Ohman, "Karolinska directed emotional faces [database of standardized facial images]," *Psychology Section, Department of Clinical Neuroscience, Karolinska Hospital, S-171*, vol. 76, 1998.
- [141] T. Kanade, J. F. Cohn, and Y. Tian, "Comprehensive database for facial expression analysis," in *Proceedings Fourth IEEE International Conference on Automatic Face and Gesture Recognition*, 2000, pp. 46–53.
- [142] M. Pantic, M. Valstar, R. Rademaker, and L. Maat, "Web-based database for facial expression analysis," in *2005 IEEE international conference on multimedia and Expo*, 2005.
- [143] N. Sornkarn and T. Nanayakkara, "Can a soft robotic probe use stiffness control like a human finger to improve efficacy of haptic perception?" *IEEE transactions on haptics*, vol. 10, no. 2, pp. 183–195, 2016.
- [144] J. Konstantinova, M. Li, G. Mehra, P. Dasgupta, K. Althoefer, and T. Nanayakkara, "Behavioral characteristics of manual palpation to localize hard nodules in soft tissues," *IEEE Transactions on Biomedical Engineering*, vol. 61, no. 6, pp. 1651–1659, 2014.
- [145] J. Konstantinova, G. Cotugno, P. Dasgupta, K. Althoefer, and T. Nanayakkara, "Palpation force modulation strategies to identify hard regions in soft tissue organs," *PLOS ONE*, vol. 12, no. 2, pp. 1–24, 02 2017.
- [146] R. Birdwhistell, *Kinesics and context: essays on body motion communication*, 1971.
- [147] E. Cambria, N. Howard, J. Hsu, and A. Hussain, "Sentic blending: Scalable multimodal fusion for the continuous interpretation of semantics and sentics," in *2013 IEEE symposium on computational intelligence for human-like intelligence (CIHLI)*. IEEE, 2013, pp. 108–117.
- [148] I. Chaturvedi, R. Satapathy, S. Cavallari, and E. Cambria, "Fuzzy commonsense reasoning for multimodal sentiment analysis," *Pattern Recognition Letters*, vol. 125, pp. 264–270, 2019.
- [149] T. C. Yu, N. C. Wilson, P. P. Singh, D. P. Lemanu, S. J. Hawken, and A. G. Hill, "Medical students-as-teachers: a systematic review of peer-assisted teaching during medical school," *Advances in medical education and practice*, vol. 2, p. 157, 2011.

...