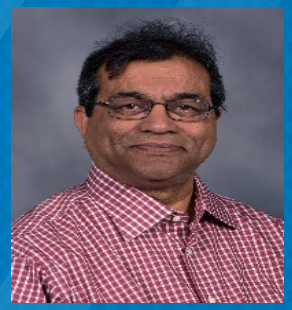
The Tenth International Conference on Intelligent Systems and Applications,
INTELLI 2021, July 18 to July 22, 2021, Nice, France,
https://www.iaria.org/conferences2021/INTELLI21.html

Keynote Speech

Current Research Directions in Social Robotics to Substitute Human Empathy in Elderly and Patient Care



Arvind Bansal
Professor, Department of Computer Science
Kent State University, Kent, OH, USA
email: arvind@cs.kent.edu

KENT STATE UNIVERSITY

Motivation and Acceptability

- Severe anticipated shortage of healthcare / elderly care assistants due to negative population growth in developed countries
  - annual cost of pain management in USA: $600 billion (Chapman and Beach, Health Communications, vol. 35, 2019)
- Three major aspects for human acceptability of social robotics
  - functionality: assisting without dominating or giving them the feeling of over dependency
  - interaction: comprehending human-actions, intentions, emotions and unexpressed feelings such as pain, irritation, frustration and communicating empathetically
  - human-acceptable form: comparable to actual form in the real world for companionship, daily activities help, medication, guidance, communication with outside world

Outline and Acknowledgement

- Requirements / research / unaddressed issues for empathy in social robotics
  - emotion analysis; pain analysis; gesture analysis
  - empathetic haptic gesture generation
  - dialog analysis/generation
  - e-skin and haptics for empathy and social touch
  - bipeds balancing

- Acknowledge all researchers whose contributions could not be cited in this limited space presentation.

Elderly Care

- Companionship
  - interacting with them meaningfully - reading a book
- Support failing eyesight
  - keeping them connected with the world news, friends and kins
  - nonobtrusive guidance while walking
- Supporting failing muscles
  - continuous nonobtrusive monitoring such as falling / preventing falling
  - helping them feel independent of kins and caretakers
  - helping them with daily needs: bathing, cooking, feeding (if needed)
- Supporting failing memory: functional, short term, comprehension, procedural
  - helping them with medication, appointments, names of persons
  - learning and helping with daily routines such as evening walks

Patient Care

- Mental Health - cognitive impairment
- Rehabilitation such as stroke, war-related damage, accident
  - upper and lower body dysfunction
  - speech therapy
  - gait improvement
- Disability assistance
- Patient transportation
  - bed to armchair / wheelchair and back; escorting patient
- Assisting daily activities
  - supporting bed to floor; bath assisting; feeding; supported walking
  - supporting medication / vital signs monitoring
  - human activity recognition to prevent fall; helping to get up from lying and sitting posture; collision avoidance in a constrained environment

Disability Assistance

- Helping with sensory limitations of elderly / patients
  - night vision, blurry vision, hearing problems, temperature sensing problems (during bathing)
- Enhancing muscle capability
  - monitoring, predicting and warning before anticipated accident (fall)
  - muscle stabilization as in Parkinson disease during walking and feeding
  - body stabilization as in learning to walk after stroke / paraplegia
  - holding objects / lifting objects
  - transporting and lifting heavier object without relinquishing human control
- Comprehending and improving degraded dialogs of elderly
- Making suggestions and asking for permission before acting
- Learning from daily activities to fill in for memory loss caused by dementia / Alzheimer's

Functionality vs Empathy

- Current research focus is on functionality
  - human nurse = functionality + empathy
  - empathy largely ignored by AI researchers
- Functionality is in early stages despite two decades of research
  - basic interaction capability and sensor-based perception
  - limited emotion understanding or generation (not much multimodal integration)
  - limited hand / head gesture analysis; almost no meaningful co-speech human gesture generation
  - basic hand motion and grasping with limited haptics capability
- Elderly / patients need empathy
  - compassion, friendliness, feeling, reassurance, warmth, encouragement
  - not reminded that they are helpless / overly dependent
  - activities / needs to be dynamically predicted and fixed for personal physical challenges
  - polite corrections / guidance without burdening or domineering
- Multiple domain specific research need to be integrated in one human-like humanoid

Required Empathy in Nursing

- Sharing patients/ elderly experience, feelings, pain and difficulties in action
  - being one with observed using mirror neurons (sharing) and cortices (mentalizing)
- Keeping the expectations from elderly / patient simple
- Helping sufficiently but not to make elderly persons feeling belittled
- Expressing compassion
  - tolerate misbehavior / frustration / agitated behavior
  - avoid negative emotion, inciting behavior or reactions
  - showing patience, warmth and tenderness
  - act in advance to relieve anticipated difficulty, pain / distress
- Encouraging dialogs: multimodal affective prosody in response to pain
  - facial expression + speech modulation (tone variation) + choice of positive words (semantics)
  - haptic social touch along with expressing kindness verbally
  - finding creative solutions to the elderly / patients' problems/agitations

Required Empathy in Communication

- Empathy in dialog: listening → comprehending → feeling → speaking / reassurance
  - some require just listening while others want solution / advise
- Speaking to express empathy: caring, tenderness, warmth and reassurance
  - frank, honest and simple with embedded explanations
  - patience against anger, disbelief, moaning, crying, aggression and denial of reality during dialog
  - overcoming hesitation to share
- Listening is as important as speaking
  - listening is both verbal and nonverbal
  - comprehending underlying conditions from expressed behavior / emotion (displaced anxiety / pain)
  - filtering perceptual distortions due to elderly memory loss or altered reality
  - situation assessment
- Dialog analysis has three major aspects
  - technical analysis: syntactic parsing and semantics analysis
  - context-based comprehension of ill-explained situation / symptoms
  - relating to the patients' underlying conditions from past knowledge (requires reinforced learning)

Current Systems vs. Human Nurse

- Current systems
  - good vision, speech recognition / generation, lower limb balancing / walking (Nao / Atlas)
  - limited NLP, gesture generation / hand grasp: (Nao, Erica, Sophia, and now Grace – COVID 19 nurse, etc.)
- Limitations of current systems
  - limiting DOF and lack of graceful human-like motion
  - high energy consumption by electromechanical / hydraulic systems
  - limited distributed skin-sensory perception to attain human-like haptics
  - limited interaction functionality: lack of emotion / pain comprehension, gesture analysis/ comprehension, behavior analysis; limited dialog comprehension / generation
  - more focus on developing domain specific functionality very far from human empathy
  - limited prediction/anticipation capability such as instant support for avoiding patients' fall
  - limited situation assessment / perceptional reality of agitated elderly person & patients
- No research in empathetical behavior generation (patience, tolerance, accommodation)

Criteria for Acceptance

- No external superiority in function. It might overwhelm elderly / patients
  - no complex learning to deal with social robot (it has to be the other way round)
  - no additional probing of body than a human nurse would do
- Same amount of care, understanding and tenderness as a human nurse
  - provide emotional safety
- Haptic medical capability
  - injections, measuring / checking vital signs, provide medication, gently lift a patient
  - assessing and predicting pain by facial expressions & paralinguistic vocalization, muscle tension
- Dialog and situational awareness capability along with reassurance capability
- Assist to simulate the same real-world interaction with the surrounding
- Letting the elderly / patient have the feeling of control of the situation

Social Robotic Healthcare Assistant

Vision (face detection)
Hand motion
Speech recognition
Speech generation
Limited conversation, vitals monitoring, scene and dialogs analysis

Image credit: REUTERS, June 2021

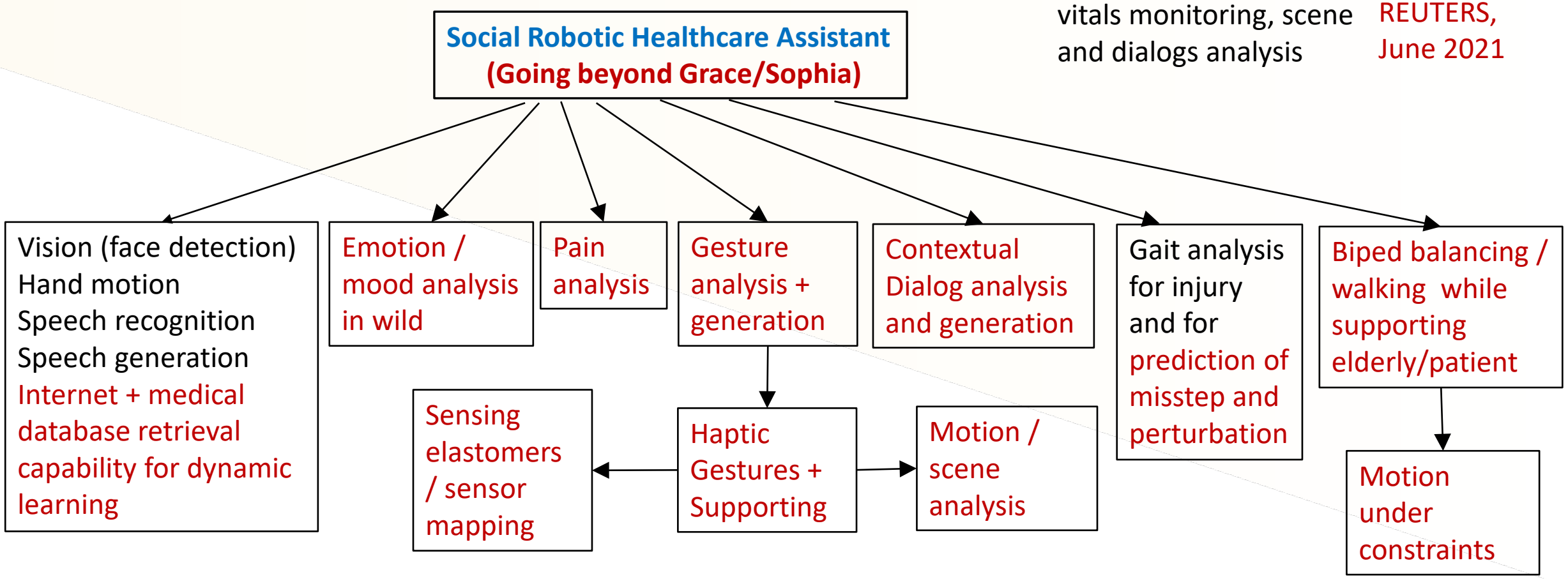
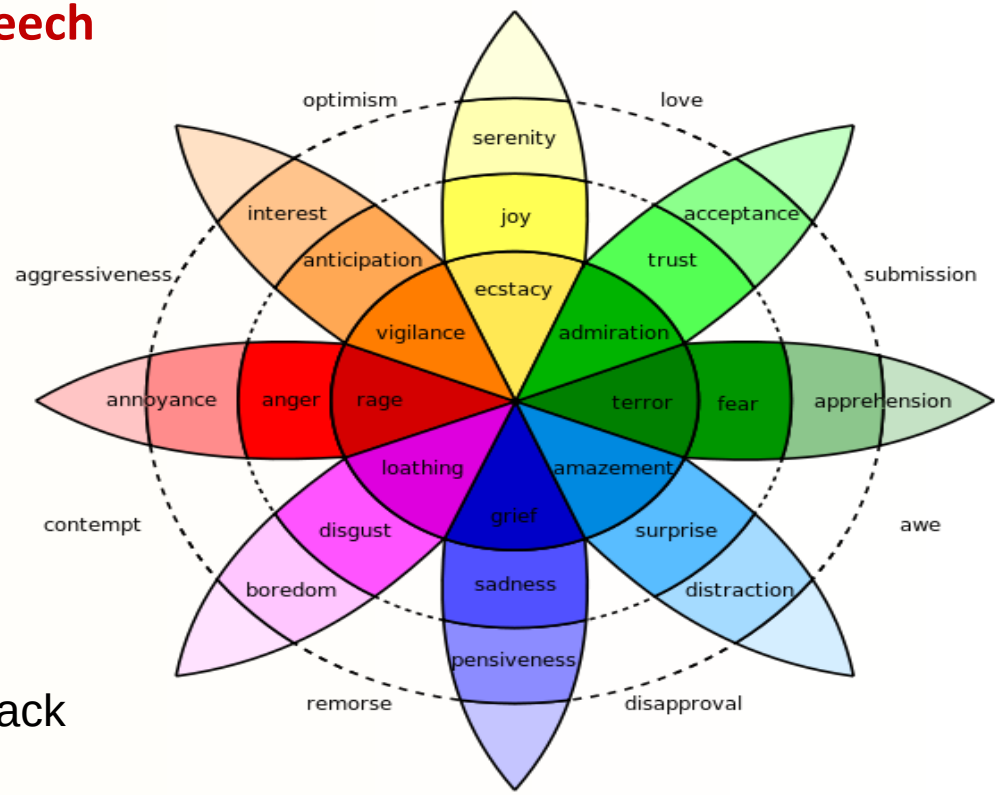


Image of Grace taken from Photos: Meet Grace, the healthcare robot COVID-19 created | News-photos – Gulf News for illustration, no copyright infringement intended

KENT STATE UNIVERSITY

Emotion Expression



Plutchik's wheel of emotion

- Facilitates interaction with the world
  - involves temporal lobe, stimulus to Thalamus regulated by cortex based on social conditioning, aging, and bodily feedback
- Classification
  - Ekman's: fear, happiness, sadness, disgust, surprise, anger (Ekman and Friesen, 1977)
  - Plutchik's wheel of emotion: Ekman's classification + trust and anticipation based on valence and arousal; mixing of basic emotions. (Plutchik, 1980)
- Missing mental states
  - anxiety, frustration, irritation, impatience, pain, depression, abruptness, confused
  - needs multimodal integration for analysis: dialogs, facial expressions, speech modulation, gesture, tears, silence, paralinguistic vocalization, and behavior pattern

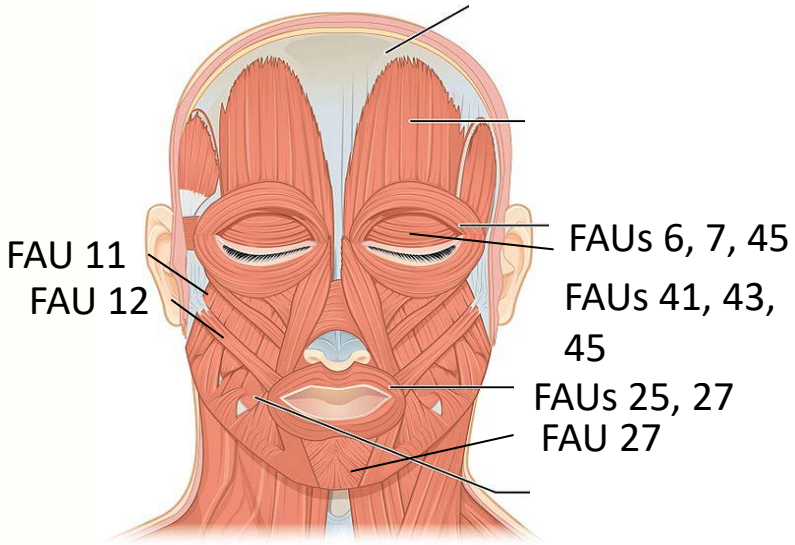
Image taken from Wikimedia Commons, public domain

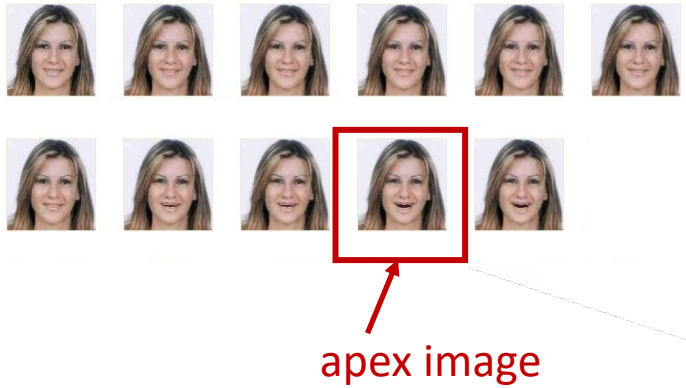
Citations: 1. Ekman and Friesen, "Nonverbal Behavior," Comm. and Soc. Interaction, 1977
2. Plutchik, "Emotion: A Psychoevolutionary Synthesis," Harper & Row, 1980

KENT STATE UNIVERSITY

Facial Expression Analysis

- Limited to Ekman's six basic facial expressions
  - disgust, fear, happiness, sadness, surprise,
  - mapping facial expressions to FAU (Facial Action Units)
  - AI classification techniques such as SVM, CNN, HMM, Bayesian classifiers and their variants
- CNN variants improve recall by 10 – 15% over FAU models using curated databases and static images
- Real life involves skewed head, head rotation and occlusion
  - occlusion by external objects, shadows, poor lighting, conversational gestures and multi-party interaction
  - modeling moving / rotating head
  - CNN is not rotation invariant
  - apex facial expression image selection during video analysis





Real Life Scenarios (Nurses' Views)

- Images are rotated, skewed, partially occluded due to rotation, lighting conditions, other people and are in continuous motion during conversation
- Imagine the nurses' views in the pictures. They have to assess pain and facial expressions using skewed and rotated face.

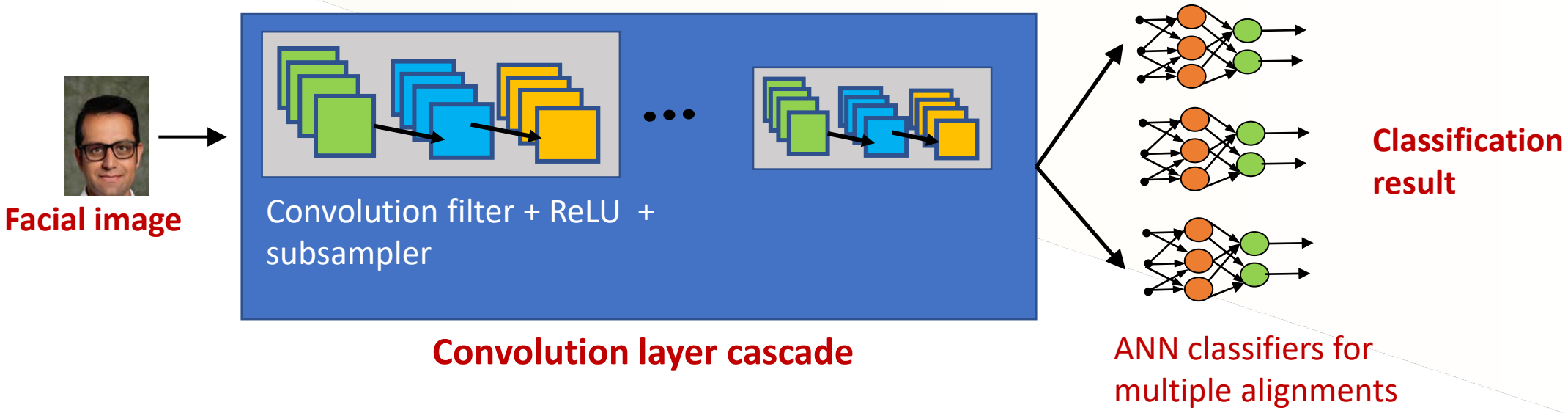




Images taken from Wikimedia commons public domain

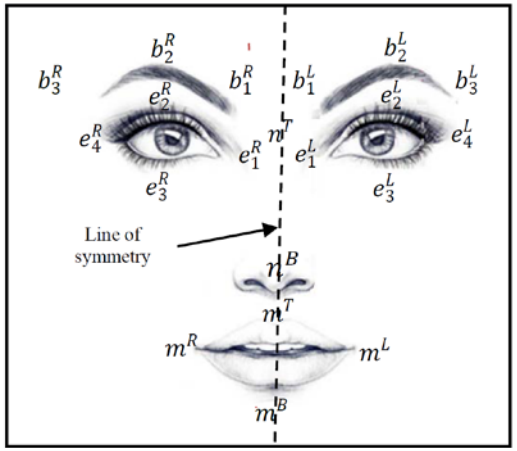
Multiple Alignment CNN models

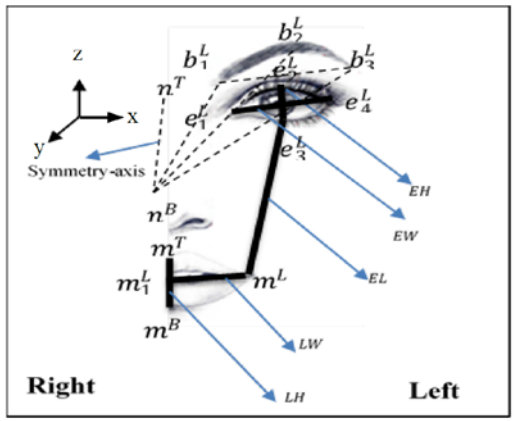
- Two types of layers
  - cascade of convolution filters
  - multiple second stage ANNs for classification under different alignments



Handling Motion + Occlusion

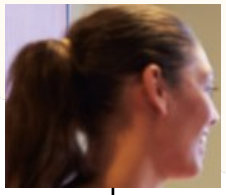
- Problem
  - feature-points are occluded with deviation from frontal pose causing performance loss
- Proposed solutions by researchers
  - texture averaging from global and other local patches
  - symmetry + attention to recover occluded feature-points
  - multiple alignments CNNs: one ANN for each rotational alignment
  - combination of multiple alignments CNN, Long Short-term Memory (LSTM) and transfer learning to find out the best mapping of the rotated or skewed postures (Li et al., IEEE Access, vol. 7, 2019)
  - Integrate multiple alignment CNN + FACS symmetry based geometric modeling (Ghayoumi and Bansal, Intelli 2021)





Images credit: Bansal and Ghayoumi, Intelli 2021
Citation: Zhang et al., "Facial Expression Analysis Under Partial Occlusion: A Survey," ACM Computing Surveys, vol. 51, Article 25, 2018

CNN + FACS for Handling Head Rotation



Denoising

Symmetry-based Geometric Model Classifier

Feature-point extractor → Invariant ratio & angle calculation → Symmetry-based FAU movement calculation → FAU based classifier

Angle value stream

Angle value stream

result selector

Multiple alignment CNN Classifier

Cascade of convolution layers

Convolution +ReLu + subsamp. layers 1 → Convolution + ReLu + subsamp. layers 2 → Convolution + ReLu + subsamp. layer 3

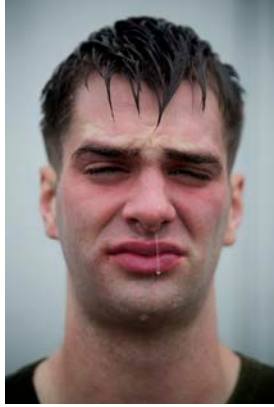
ANN selector

multiple ANN – one for each alignment

ANN result integration

Types of Pain

- Acute pain caused by recent rupture, intrusion, penetration of tissue
  - verbally and / or nonverbally by detecting symptoms of distress
  - tissue texture and color analysis
- Sudden physiological changes / increase in pain intensity can be detected
  - sweat increases skin conductance; change in skin color
- Chronic pain in joints, shoulder / muscular / surgical / lower back / fibromyalgia
  - not expressed nonverbally unless the intensity becomes high
  - expressed as mood disorder, depression or anger
- Transient pain
  - injection, sudden stress to tissues
  - pain during transferring, ambulation, and repositioning



Pain Expression and Detection

- Verbal
  - paralinguistic vocalizations: pain cries such as arghh*hhh(*hm)" , sobbing, screaming, moaning, gasping, groaning, grunting, whimpering, noisy breathing (Herman, Weijelberg et. al., 2020, EJP)
  - modulated speech associated with paralinguistic vocalization / tears
- Nonverbal
  - gestures such as holding body parts and involuntary grimace with facial-expressions
  - physiological activity, such as pallor, flushing, sweating, muscle tension
  - altered action or lack of regular actions such as limping or not fully stretching a limb
  - trauma signs such as excessive bleeding or bruise signs
  - temperature and color change of bruised / inflamed area using Infrared image analysis
  - contextual information / scene analysis such as presence of sharp object, fall silhouette detection
- Pain assessment in dementia is a challenge due to limited communication
- Chronic pain gets displaced and expressed differently such as depression and mood disorder

Facial Muscles and FAUs in Acute Pain

- Brows lowerer (FAU 4); Cheek raiser (FAU 6)
- Eye-lids tightener (FAU 7); Nose wrinkler (FAU 9)
- Nasolabial deepener (FAU 11);
- Upper lip raiser (FAU 10); Lip corner puller (FAU 12)
- Lip stretcher (FAU 20); Lips part (FAU 25)
- Jaw drop (FAU 26); eye-lids droop (FAU 41)
- Eyes closed (FAU 43); blink (FAU 45)

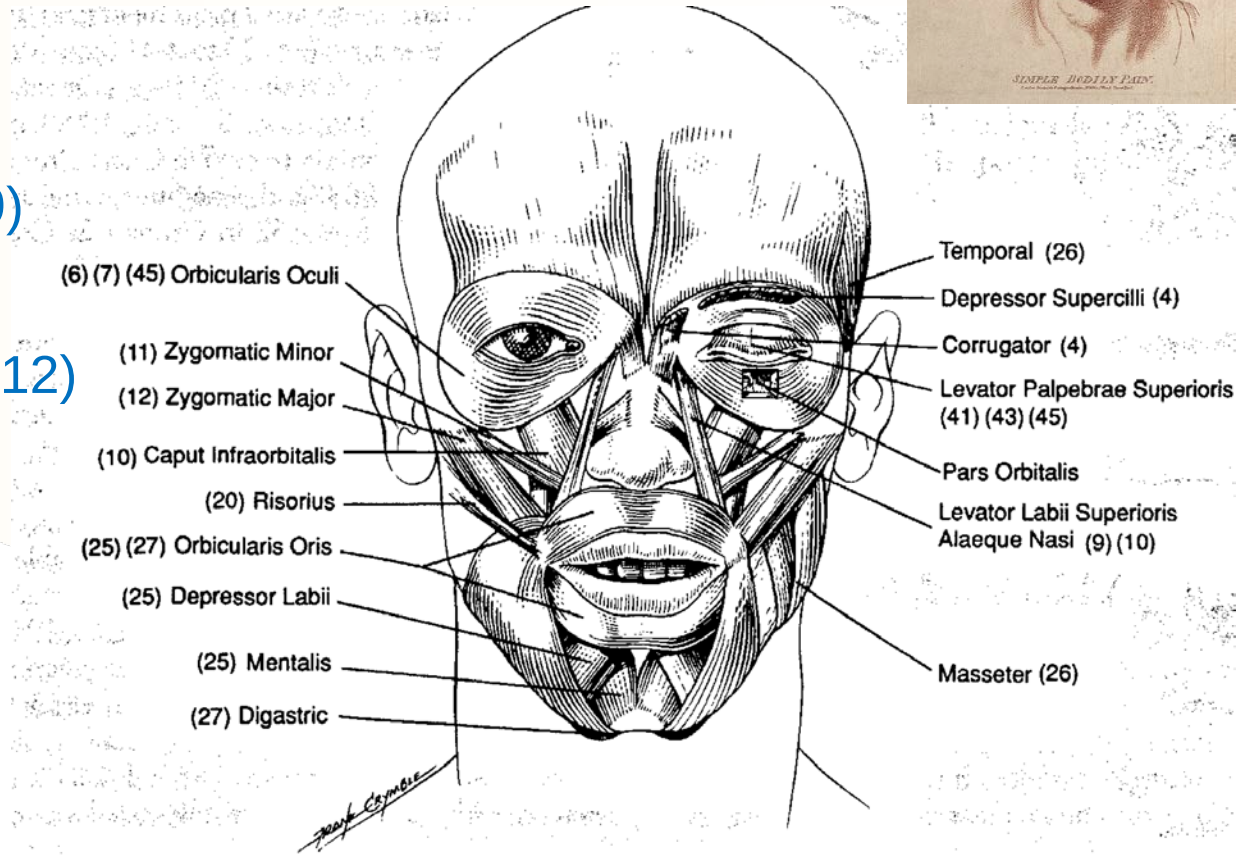
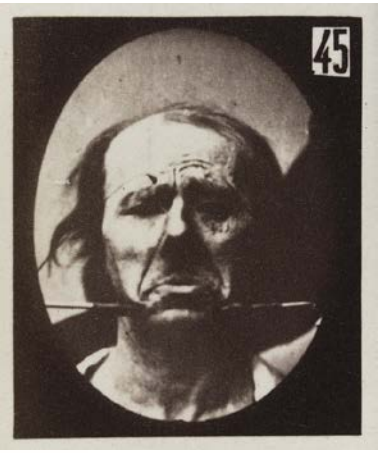


Image taken with written permission from Professors Craig and Prkachin
Citation: KD Craig, KM Prkachin, RE Grunau, The facial expression of pain., In D. C.Turk & R. Melzack (Eds.), *Handbook of pain assessment* (pp. 117–133). The Guilford Press, 2011

KENT STATE UNIVERSITY

Automated Analysis of Acute Pain



- Physiological methods: Epidermal conductance increase during pain (EDA)
  - (Gunther et al. 2016, Acta Anaesth. Scandinavica)
- Computational classification techniques of facial expression analysis
  - SVM, mixed Gaussian, CNN + transfer learning, and FAU analysis
  - Multiple datasets: CK+, UNBC-McMaster and Willkie's cancer patients' dataset
  - Transfer learning handles variations in pain dataset belonging to different subjects
- Multimodal integration: EDA + Facial expression
  - (Susam et el., EMBC 2018)
- Issues in facial expression analysis
  - overlapping FAUs for facial expressions and pain
  - face may be skewed or partially occluded
  - intensity is subjectively expressed
  - chronic pain gets displaced and not expressed directly
- No research on signal processing of paralinguistic vocalization associated with pain

Health-related Haptic Gestures

- Function related
  - pressing arm for injection / drawing blood
  - pressing body parts to test for pain
  - cleaning wound and applying bandage
  - supporting the patient to stand/walk
- Empathy related
  - wrapping arm around patient to provide support / reassurance
  - holding hands to show warmth or guidance
  - massaging body part for blood circulation
  - raising the patient gently from lying position to sitting position




Images taken from Wikimedia commons public domain

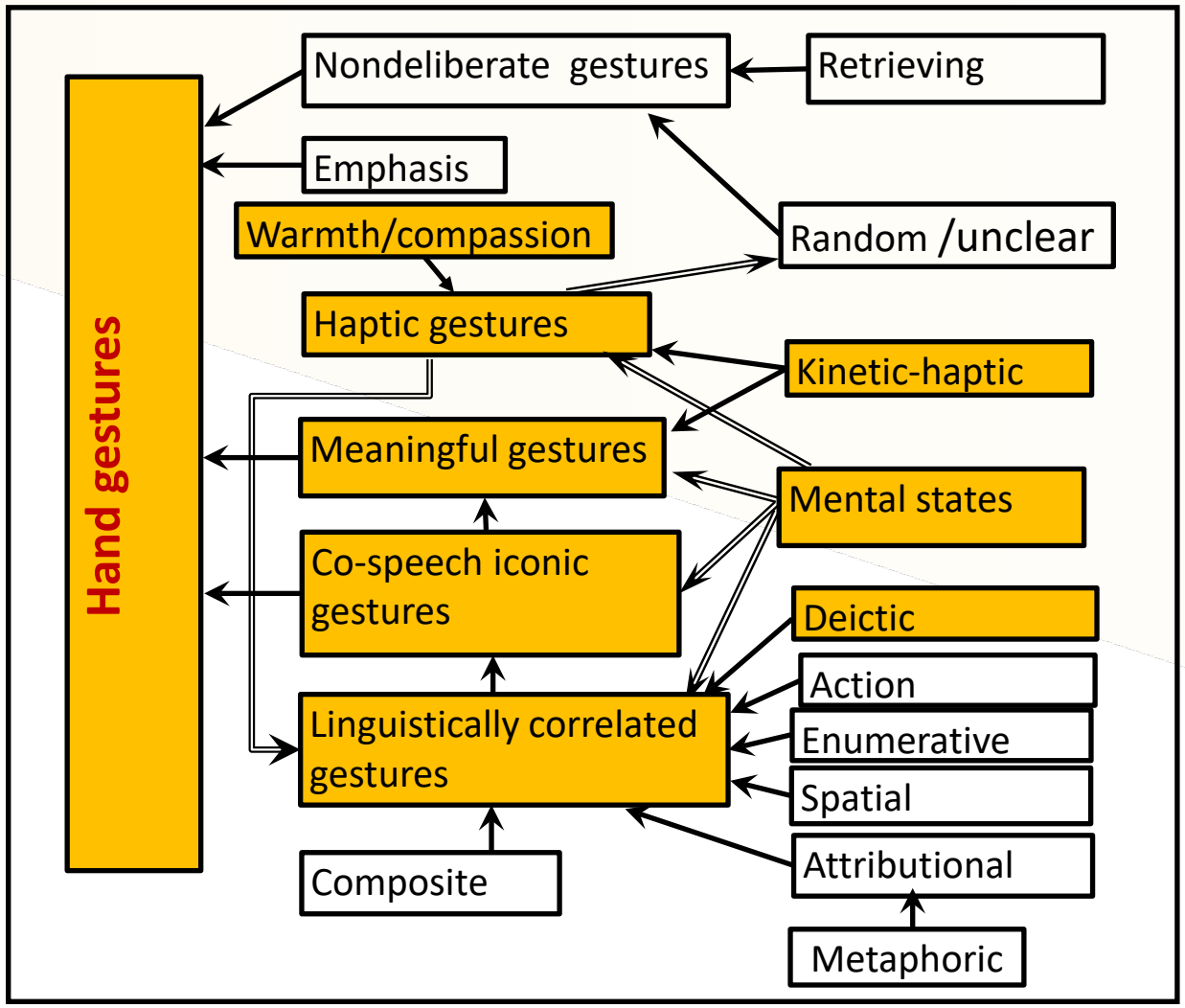
Gesture Analysis by a Human Nurse

Haptic gesture + pain analysis

- Helping with daily activities: walk / eat breakfast / take medicine / bath
  - helping in anticipation
  - analyzing support organs contact with surface, posture and gait
- Pain assessment
  - analyze hand / head gestures for pain / localization
  - muscle tightening, pain facial expressions and paralinguistic sound
  - knowledge of other symptoms and abnormal gait / human activity analysis
  - localization of pain by applying measured pressure
- Situation assessment
  - depression may be expressed by posture
  - comprehending silent gesture for request to help
  - anxiety assessment: fist tightening, clasped hands, gaze avoidance
- Question: where does it hurt?
- Reassurance: Honey, It will just pinch a little!

Image taken from Wikimedia Commons – public domain

Hand Gesture Classification for Empathetic Healthcare



single parent subclass

multi parents subclass

Health related gestures

- Deictic gestures are used to point to an object in the scene

Image adopted from Singh and Bansal, Computing Conference, London, UK, July 2021

Required Haptics for Social Touch

- Social touch
  - force sensing, pressure sensing, vibration sensing, temperature sensing
  - fingers, palm: pressure sensing, temperature sensing, vibration sensing (counting pulse)
  - arm: force sensing for providing support to patient / elderly
  - flexible sensors embedded in elastomers
- Perception and monitoring of localized sensors in real-time
  - sensor feedback and integrated analysis with vision and sound
- Movements of micromotors groups to express different social touch
- Motion analysis to predict anticipated fall using force sensor in arms / vision
  - flexible rigidity to provide just enough support
- Motion analysis in proximity
  - posture / gait perturbation analysis to predict imbalance or failure to change posture

Required Research for Soft Haptic Hand

- Flexible e-skin related research
  - elastomers (hydrophobic silicone polymers) to provide thin sensory deformation resistant e-skin
  - flexible micro and nano sensors embedded in e-skin for optimum distributed force: piezoelectric, resistive, capacitive, triboelectric
  - mapping sensors for optimum sensory perception
  - soft hand gripping and grasping for applying optimum force
  - self monitoring for damage detection, nanostructured pathways for feedback and self-repair

  (Khatib, Zohar et. al. Advanced Materials, 2020)
- Movement of hands in a constrained environment
  - improving degree of freedom specially fingers and thumb movement / coordination
  - collision avoidance in constrained environment
  - avoiding entanglement with wearable devices
- Fast alignment with optimum human-like haptic force is missing

Gesture Analysis/ Generation Research

- Limited classification of hand-gestures in psychological domain
  - haptics gestures and pain gesture classification still not addressed
- Gesture analysis is limited to a subset of hand-gesture and head gestures
  - synchrony in head-gestures recognition is just being addressed (Singh and Bansal, FTC 2021)
  - Multiple AI techniques for hand gesture analysis: CNN + LSTM, DBN, HMM, depth silhouette
- Limited head-gesture / hand gestures generation in humanoids
  - limited grasping and arm motion arthritic gait analysis of pelvic / knee joints
  - DOF unsuitable for haptics operations such as injection / wound cleaning / supporting a patient or testing pain
  - lacking social touch capability to express warmth / reassurance or micro-pressure sensing capability for injection, drawing blood, etc.
  - 36 head gestures with limited speech demonstrated by Singh and Bansal (2019)
- Pain-related gestures analysis with paralinguistic vocalization still not addressed

Required Dialog Analysis

- Patient's dialog may be disoriented due to pain / agony / personal beliefs / lack of medical knowledge / social withdrawal or their combinations
- Varying time to explain the same phenomenon by different patients
- Dialog analysis requires
  - using ontology to translate patients' description to medical terms and vice versa
  - consistency verification using verified medical knowledge base
  - removal of the noise caused by patient's belief/ perception / knowledge / explanation
  - summarization
  - assessment of the patient's mental state / mood for filtering out displaced behavior
  - corpus analysis to detect words and phrases related to anxiety / pain / agony / frustration
  - integration with speech modulation analysis to detect anxiety / pain / agony / frustration
  - picking distress (such as cancer patients) related cues and sensitivities from multimodal analysis – requires learning patients' traits
  - recording electronically into patients' database using standardized database format

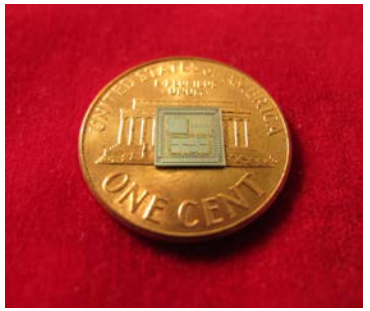
Required Empathy in Dialog Generation

- Silence and tenderness – not every patient's dialog needs a speech response
  - use of age / gender / culture specific reassuring words
  - ensuring that patient feels being tended to
  - substituting words by haptics for reassurance: holding hands / arms around shoulders
- Explanation capability embedded in dialog generation
  - large percentage of patients are hesitant to ask questions
- Dialog generation requires
  - ontology + knowledge base of patients' knowledge to explain related medical knowledge
  - filtering out information that may arouse anxiety / agony
  - summarization with embedded explanation
  - active medical information sharing, including information on computer screen
  - integration with co-speech haptic gestures for reassurance
  - accessing Internet and medical database like Alexa to answer patients' simple queries
  - using encouraging positive words for participation / reassurance mixed with listening

Dialog Analysis Research

- Dialog analysis
  - frame based semantic studies
  - limited co-speech hand-gesture and dialog alignment studies (Wagner et el, Speech Comm., 2014)
  - modeling dialogs using conceptual dependency and predicate logic
  - anaphora resolution (pronoun disambiguation)
  - summarization and topic extraction using latent semantic analysis
- Recent research for empathy in dialog analysis
  - LSTM and attention based dialog generation to maintain context
  - emotion expressive encoder-decoder: (utterance, speaker) ► emotion + GAN (Generational Adversarial Network) / reinforcement learning
- Issues largely unaddressed
  - personality factors in elderly / patients: Introvert/extrovert; friendliness; conscientiousness; emotional stability; openness
  - behavior alterations with emotional state, anxiety, anticipation, and displaced pain

Biped Balancing for Social Robots



Timing and IMU

- Current technology: Atlas (Boston Dynamics) and/or Nao (CMU)
  - general balancing during walking: IMUs to measure pitch, roll and yaw; inverted pendulum
  - external continuous force resistant balancing (Kim and Lee, Robotica, 2015)
  - real-time balancing on slippery / uneven surface (yaw control, CoM/ ZMP, foot rotation measurement, change in angular momentum, measuring gait perturbations)
  - balancing while picking /grasping small objects (combining ankle movement + hip pushed backward + stepping forward strategy (Shen et al., IEEE access 2020)
  - fall recovery by pushing back and joint movements
- Needs significant improvement
  - adaptive balancing / walking while supporting patients (sudden & persistent external force)
  - human-like elegance during walking – emulating human pelvis (Ohnuma, Lee and Chang, Intel Serv Robotics, 2017) – adaptive COG, nonlinear tilt and rotation; currently excessive torque on knee / ankle actuators
  - medical scene analysis and object recognition to avoid small obstacles in a constrained environment like living room or wearable equipment around patient
  - long-time lightweight power support with high load / robot weight ratio

Conclusion

- There is a need to incorporate empathy related research in social robotics
  - automated pain / emotion / behavior analysis
  - prediction using gait analysis / activity perturbations
  - sensor-based elastomer skin and its mapping for localization and sensory perception
  - movements of micro actuators for haptic gesture generation in social touch
  - elegant human movements in a constrained environment
  - indirect cues analysis for distress in vocalization and dialog analysis
  - compassionate behavior generation