# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>3</td>
</tr>
<tr>
<td>History</td>
<td>3</td>
</tr>
<tr>
<td>Implications to Embedded Applications</td>
<td>4</td>
</tr>
<tr>
<td>Perceptual Computing Algorithms are Turbo-charged by GPU Compute</td>
<td>5</td>
</tr>
<tr>
<td>AMD is a Leader in Perceptual Computing</td>
<td>6</td>
</tr>
</tbody>
</table>
Perceptual computing promises a more natural way to interact with machines. Typically the term is applied to personal computer features such as facial recognition, used to replace passwords, or gesture analysis, used as an alternative to input from a computer mouse. However there are recent developments in software technology made practical by graphics processing unit (GPU) compute that are expanding the concept of perceptual computing, creating a drastic uptick in the functionality of embedded applications.

To perceive is to assign meaning to what we see, hear, touch, taste and smell. It’s second nature to us humans, but extremely difficult for a computer. It’s not like scientists have not tried; they have, in fact, been working on this problem for decades.

History
In 1936 Alan Turing famously described a computing machine that could input and process symbols in such a way as to simulate any type of computation. The Turing test pits a computing machine against a human. As the experiment is described, given the ideal Turing Machine, an expert would not be able to tell the difference between the response of the computing machine and the response of a human when posed with the same question.

Actually implementing such a machine has evaded computer scientists for decades. But this is finally changing. Perceptual computing is opening new channels of communication from man to machine. Machines, in turn are beginning to respond back in humanoid ways. We already have multiple examples of applications that simulate intelligent human/machine interaction using natural language.

This is exciting enough, but the development that will be a revolution for the interaction of man and machine is the capability of a machine to interpret non-verbal cues from secondary modes of human communication such as gestures, pose (body language), vocal intonation and eye/facial expressive nuances. These cues are often more definitive of the true intentions of the human psyche than mere verbal expression. This capability will fundamentally change the type of information that a computing device can extract from interaction with a human.

In a TED Women talk, computer scientist Rana el Kaliouby states, “...our human face happens to be one of the most powerful channels that we all use to communicate social and emotional states, everything from enjoyment, surprise, empathy and curiosity.”

Perceptual computing is no longer the stuff of science fiction. Currently there are commercial software products that can analyze facial muscle movements and assign affective meaning to them with a high degree of accuracy. This exciting development will naturally progress to include the capacity to extract affective meaning from human gestures, body posture and movements, eye movements, and physiological responses (e.g., sweating, and heart and respiration rate).

Marian Bartlett, a researcher at the University of California, San Diego, has developed software that identifies and tracks “micro-expressions” so brief that the person emoting them is not even aware that they are occurring. These expressions are often incongruous and so fleeting that other human observers may simply be confused about their possible meaning. However, computer analysis of such sequences of micro-expressions can reveal complex emotional reactions to external stimuli. These involuntary affective signatures cut through all pretense and social urbanity and can be used to predict future conscious decisions.

Emotional state information can be corroborated by physiological response data. Obtaining physiological data no longer requires attaching instruments to the subject. Software applications are already available to measure heart rate and respiration in a totally non-invasive manner merely by analyzing video of the face.

Even if a person can mask their true feelings and intentions with a “poker face,” physiological data can indicate that a person is lying or withholding information. If the person is using a smile to mask their true feelings, action unit analysis software can distinguish the difference between a genuine smile and a forced smile.
New developments in human/machine interaction do not stop with discerning the emotional state or intentions of humans. Classification software can accurately gauge the gender, age, and ethnicity of the person on the other side of the communication exchange based on facial features. This information can be used to tailor the content and style of the interaction to the stereotypical preferences of the demographic group to which the person (apparently) belongs.

Given the vast amount of tagged image data available on the Internet and the ubiquity of social media, it may be possible to identify the individual and use personal data to target the information delivered based on the person’s expressed interests.

**Implications to Embedded Applications**

The implications of these radical new developments to embedded computing devices are endless. For example, with these capabilities, a digital signage device could observe the viewer and determine the demographic group of that person. Furthermore, it could determine the general emotional state of the person and the level of engagement with the content being displayed allowing the device to dynamically change the content and style of interaction based on measured emotional responses and/or changes in the level of engagement.

If an Internet search hits on a tagged image of the viewer’s face, the device could offer the person products or services similar to those that have been purchased by this person in the past. Dynamically generating discount coupons for products of interest could be an effective marketing tool.

Add synthesized voice response and the digital signage device could completely replace, for example, the typical greeter at the big-box home improvement store. As a customer stares at the bewildering store layout schematic, the device would ask the person, “How may I help you?” Human: “Where can I find a toilet bowl auger?” Machine: “Aisle 12, half way down on the right. It should be at eye level.” And do so with a pleasant, patient and highly empathetic synthesized voice.

How about aircraft safety? Wouldn’t it be useful to know not only the level of alertness and engagement of the pilot but also their emotional/mental state? If the intentions of the pilot are known not to be in the best interests of the safety of the passengers, remote control systems could take over and land the plane, saving hundreds, perhaps thousands, of lives.

Consider the applications of perceptual computing in the area of medical devices. New video processing techniques provide a totally non-invasive/non-touch means to measure respiration and heart rate. Other apps use the flash and front facing camera of a smartphone to measure pulse rate and oxygen saturation.

Healthcare providers are increasingly working with game designers to create therapeutic games to treat post-traumatic stress disorder (PTSD), attention deficit/hyperactivity disorder (ADHD), and other behavioral and emotional disorders. Many experts are anticipating a day when games might be prescribed in place of pharmaceuticals for disorders like depression and anxiety. Games might even become an effective tool for the diagnosis of psychiatric disorders.

Autistic persons find it very difficult to understand the non-verbal social and emotional cues of others. A simple augmentation device alerting them to these cues could help them interact with others in more socially acceptable ways.

Bio-feedback can help us relax and reach a calm, focused meditative state. In the past bio-feedback has only been available in a doctor’s office. What if everyone had access to this therapeutic technique? A company currently offers a brainwave sensing device that attaches wirelessly to a smartphone or tablet and includes a meditation assistance app.

Then there is human/machine interface (HMI) for industrial control. Imagine an industrial control scenario where the device is capable of inferring operator emotional state information from facial expressions and eye movements.
Machine to human operator: “You seem concerned about the pressure in reaction vessel number 12. I suggest you reduce the flow rate of reactants by 3 percent and decrease the temperature of the vessel by 10 degrees centigrade.” Human: “Sounds like a good idea. Please make the suggested adjustments.”

In infotainment systems the implications are obvious: by watching the reaction of the human recipient to the current content stream, the device can tailor the content to the emotional state and preferences of the person. Automobile infotainment systems could be integrated with driver alertness monitoring devices so that the content being delivered could actually positively influence the driver’s alertness and engagement in the driving process.

Perceptual computing informs and provides new insights into the processing involved in making natural, immersive telepresence systems. These systems require the processing of time-aligned video and audio data streams. Image processing/facial recognition software could identify which person in a conference room is the current speaker and digital signal processing utilizing an array microphone could focus the sensitivity of the audio pickup to capture and clarify what the current speaker is saying. Lip-shape processing could even be used to clarify indistinct words in the speaker’s oral presentation.

Consider how perceptual computing transforms gaming: Imagine a virtual world where the computer generated (CG) characters the player meets seem to understand what the player is feeling and react in empathetic ways. Game software could determine what the player is feeling based on feedback received from the player’s face. To complete the illusion, the characters that inhabit the virtual world could themselves express emotions in a believable way. Advances in photorealistic rendering are making it possible to synthesize CG characters with emotionally expressive faces.

Measuring the player’s response to the game play places the emotional/physiological response of the player in the middle of the control loop that drives game play forward. Emotionally intelligent interaction between the game and the player opens up new gaming scenarios and could create a whole new genre of virtual reality games.

Computing devices with perceptual capabilities allow machines to interact with humans using human modes of communication rather than forcing the human to adapt to the personal computer paradigm of navigating through multiple levels of screens and menus to find the right command or setting required to carry out the desired task or obtaining the desired information. This eliminates the learning curve users experience when interacting with an unfamiliar device.

Research has shown that test subjects report higher levels of social presence (the subjective experience of mutual understanding and emotional connection/closeness) when interacting with an emotionally expressive computer-generated character as opposed to a non-expressive one. Studies suggest that this experience is impacted more by the realism of the affective/social cues provided by the computer-generated character than the belief that the character was an avatar (controlled by another human) or an agent (entirely controlled by a computer). As the visual fidelity of computer-generated agents increases to the level of photorealism, users will begin to accept the agent as an advisor, counselor, colleague and companion.

Perceptual Computing Algorithms are Turbo-charged by GPU Compute

Algorithms used to process visual and auditory input are highly compute-intensive. Fortunately they can typically be formulated in ways that allow the simultaneous execution of thousands of threads of the same code operating in parallel on different parts of the input data stream. Data can be subdivided spatially (image data), or temporally (auditory data), or both (video streams). GPUs, which are designed to take many simultaneous streams of draw instructions and assemble an image in a frame buffer, can also be used to perform mathematical and logical operations on large amounts of data in parallel.

Perception involves categorizing input streams of visual and auditory data, extracting high-level
concepts, combining these with a knowledge base, and using this information to react, formulate plans, solve problems or carry out tasks. Responses must obey physical realities and rules of social interaction.

The human brain is extremely adept at doing all this. The type of processing that goes on in the brain, however, is fundamentally different than the numerical and logical computations that take place in a computer. The brain is organized as a mesh of 100 billion or so neurons interconnected by thousands of billions of synaptic connections.

Scientists have unraveled much of the inner workings of this neural network. Neuron firing rate indicates the strength of response of a neuron to the weighted average of the activity of the pre-synaptic neurons to which it connects. Some synaptic connections inhibit the rate of firing and others increase it. Mathematical models of this biological form of processing confirm that neural networks are very powerful classifiers/pattern recognizers. Thousands of associations made between what the brain has learned and recognizes with current external and internal stimuli allow the brain to make sense of the world and react to it.

GPUs are well-suited for simulating neural networks. The firing strength of a neuron can be modeled by a floating point number and excitatory and inhibitory synaptic connections can be modeled by positive and negative weights (also represented using floating point numbers). The simulation is carried out in discrete time increments. The current output state (firing rate) of each simulated neuron in the network is evaluated using a summation of the product of the presynaptic firing strength and the associated excitatory/inhibitory strength of the connection.

To complete the process of intelligent interaction with humans requires the computer to carry out decision-making and planning. Planning and decision-making requires a schema (systematized body of knowledge) and a goal (why are we having this conversation?). Algorithms for these types of processing require searching and processing large amounts of data. GPU compute can be used to perform these types of operations as well.

**AMD is a Leader in Perceptual Computing**

AMD is a leader in providing foundational hardware and software technologies for perceptual computing. The Heterogeneous System Architecture HSAIL language enables seamless programming across CPUs and GPUs. AMD’s support of open standards for parallel computing such as OpenCL™, OpenCV and Open VX enables speedier development of perceptual computing software.

AMD Embedded Solutions (AES) offers a wide range of hardware products well-suited for implementing perceptual computing applications. AMD accelerated processing units (APUs) integrate the traditional CPU and GPU on a single die. Discrete AMD Radeon™ graphics products and board-level discrete Radeon graphics cards have enough power to handle products that meld 2D/3D graphics with perceptual computing algorithms.

The AMD Embedded R-Series SOC (formerly codenamed “Merlin Falcon”) debuts our most capable embedded product to date. With up to four “Excavator” x86 cores, integrated AMD Radeon™ HD 10000 graphics (based on AMD’s 3rd generation GCN architecture) and full HSA support, the AMD R-Series SOC is the ideal choice for digital signage, industrial control and automation, surveillance and communications infrastructure applications.

Configurable TDP (12W to 35W) allows the AMD R-Series SOC processor to adapt to your system packaging needs rather than the other way around. An integrated multi-lane PCIe® Gen 3 controller means lower system chip count. Dual 64-bit memory channels and support for both DDR4-2400 and DDR3-2133 DRAM provide plenty of bandwidth to system memory, and integrated multimedia coprocessors provide video encode/decode offload support for 4K media with compatibility for the most advanced video codec technologies.

For more information on the AMD Embedded R-Series SOC, see [www.amd.com/r-series](http://www.amd.com/r-series).
www.amd.com/embedded

FOOTNOTES
4) Ibid.
7) http://www.vitalsignscamera.com/; The Vital Signs Camera Application by Philips Innovation is available from the iTunes store.
13) See must FM the brain sensing headband, http://www.chronosense.com/

The information contained herein is for informational purposes only and subject to change without notice. While every precaution has been taken in the preparation of this document, it may contain technical inaccuracies, omissions and typographical errors, and AMD is under no obligation to update or otherwise correct this information. Advanced Micro Devices, Inc. makes no representations or warranties with respect to the accuracy or completeness of the contents of this document, and assumes no liability of any kind, including the implied warranties of merchantability or fitness for particular purposes, with respect to the operation or use of AMD hardware, software or other products described herein. No license, including implied or arising by estoppel, to any intellectual property rights is granted by this document. Terms and limitations applicable to the purchase or use of AMD’s products are as set forth in a signed agreement between the parties or in AMD’s Standard Terms and Conditions of Sale.

AMD, the AMD Arrow logo, AMD Radeon, and combinations thereof are trademarks of Advanced Micro Devices, Inc. Other product names used in this publication are for identification purposes only and may also be trademarks of their respective companies. PCIe is a registered trademark of PCIe-SIG Corporation. HDMI is a trademark of HDMI Licensing, LLC. Windows and DirectX are registered trademarks of Microsoft Corporation. OpenCL is a trademark of Apple Inc. used by permission by Khronos.

© 2015 Advanced Micro Devices, Inc. All rights reserved. PID# 157912-A