CHAPTER

11

Affect Measurement: A Roadmap Through Approaches, Technologies, and Data Analysis

*Javier Gonzalez-Sanchez**, *Mustafa Baydogan***, *Maria Elena Chavez-Echeagaray**, *Robert K. Atkinson**, *Winslow Burleson*†

***Arizona State University, Tempe, AZ, United States; **Bogazici University, Istanbul, Turkey; †New York University, New York, NY, United States**

Affect is inextricably related to human cognitive processes and expresses a great deal about human necessities (Picard, 1997); affect signals what matters to us and what we care about. Furthermore, affect impacts our rational decision-making and action selection (Picard, 2010). Providing computers with the capability to recognize, understand, and respond to human affective states would narrow the communication gap between the highly emotional human and the emotionally detached computer, enhancing their interactions. Computer applications in learning, health care, and entertainment stand to benefit from such capabilities.

Affect is a conceptual quantity with fuzzy boundaries and with substantial individual difference variations in expression and experience (Picard, 1997). This makes measuring affect a challenging task. This chapter does not intend to present a comprehensive survey of the methods used to measure affect but instead provides a roadmap through a selection of the principal approaches and technologies. Section "Affect, Emotions, and Measurement" introduces the concepts and background behind affect measurement. Section "Gathering Data: Approaches and Technologies"

Emotions and Affect in Human Factors and Human–Computer Interaction. http://dx.doi.org/10.1016/B978-0-12-801851-4.00011-2

presents approaches and technologies, explaining the type of data gathered from each, its characteristics, and its pros and cons. Section "Data Handling: Sampling, Filtering, and Integration" describes the stages of data handling, which include data sampling, data filtering, and data integration from a variety of sources. Section "Data Analysis" describes tools and techniques for data analysis that correlate affect measurements with stimuli, and presents examples of the application of these tools in the analysis of data samples collected in experimental studies.

AFFECT, EMOTIONS, AND MEASUREMENT

Affect is a construct of neural activity and psychological reactions; it is used as an encompassing term to describe emotion, feelings, and mood because they are so closely related and almost simultaneous in occurrence. Although emotion and mood are states of mind and, as such, are indicators of experiencing feeling or affect, the terms emotion and affect are frequently used interchangeably because they are so closely related.

Some theories propose that emotions are states embodied in the peripheral physiology and assume that a prototypal electro-physiological response exists for each emotion. Therefore, emotions can be detected by analyzing electrophysiological changes and identifying patterns associated with a particular emotion. Automatic affect recognition is a two-step process. First, data is gathered from electrophysiological manifestations of affect using sensing devices; these devices range from brain–computer interfaces (BCI), eye-tracking systems, text-based recognition, and cameras for facial gesture and body language recognition to physiological sensors that collect data regarding skin conductance, heart rate variability (HRV), and voice features, among others. Second, the vast amount of data gathered by the sensors is processed with the aim of inferring affective states by applying machine learning and data mining algorithms; commonly used machine learning and data mining algorithms include rule-based models, support-vector machines, Bayesian networks, hidden Markov models, and neural networks, as well as *k*-nearest neighbors, decision trees, and Gaussian mixture models (Calvo and D'Mello, 2010).

Using several sources of data, either to recognize a broad range of emotions or to improve the accuracy of recognizing a single emotion, is referred to as multimodal affect recognition. Multimodality requires a third step, integrating the information. This step either integrates the data from a number of sensing devices before running the inference process or integrates the inferences made from each device's data separately. An analysis of integration approaches and methods is presented in Novak et al. (2012). Fig. 11.1 summarizes the three steps in multimodal affect recognition.

The integration of inferences requires the adoption of a model that describes the relationships between those inferences, each usually

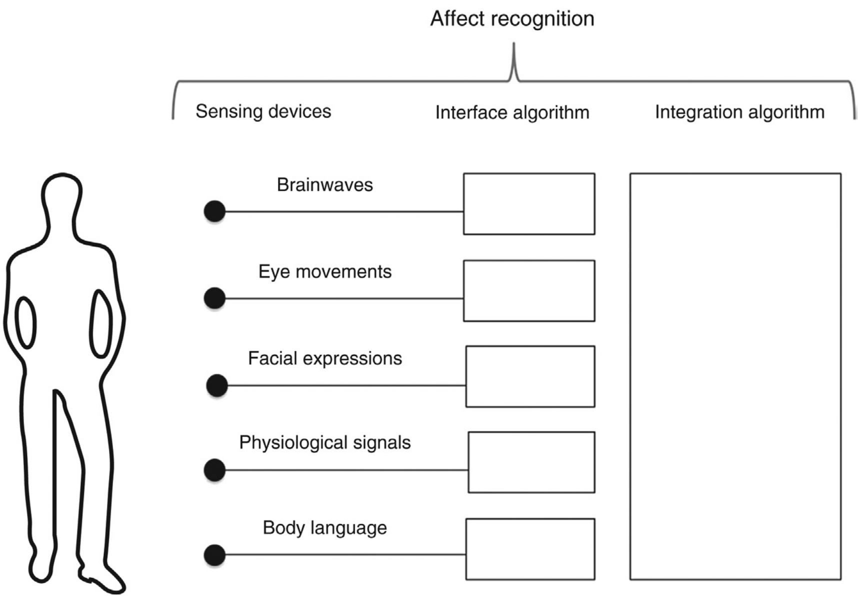

FIGURE 11.1 Multimodal affect recognition three-step process. *Source: A modified version of Fig. 1 from Gonzalez-Sanchez, J., Atkinson, R., Burleson, W., Chavez-Echeagaray, M.E., 2011. ABE: An agent-based software architecture for a multimodal emotion recognition framework. Presented at the Proceedings of the 2011 Ninth Working IEEE/IFIP Conference on Software Architecture. IEEE Computer Society, Washington, DC, USA, pp. 187–193.*

representing an individual emotion. These models are called emotional models (Gilroy et al., 2009). The classification of emotions is an ongoing aspect of affective science and experts still struggle to reconcile competing emotional models. To date, two emotional models have come to the fore: the discrete model and the continuous dimensional model.

The discrete model assumes emotions are discrete values with only a finite number of possible values, and that they are fundamentally different constructs (Ekman, 1992). A limitation of this model is that it focuses on strong emotions (such as disgust, sadness, happiness, fear, anger, and surprise) and it cannot accommodate a variety of closely related emotions or combinations of emotions.

The continuous dimensional model asserts that affective states are continuous values in one or more dimensions, and conceptualizes emotions by defining where they lie in that dimensional space. Russell (1980) proposed a two-dimensional model that links arousal with pleasure. Arousal measures intensity, or how energized or soporific one feels, ranging from calmness to excitement. Pleasure measures how pleasant or unpleasant one feels, ranging from positive to negative. For instance, while both boredom and frustration are unpleasant emotions, frustration has a higher level of arousal.

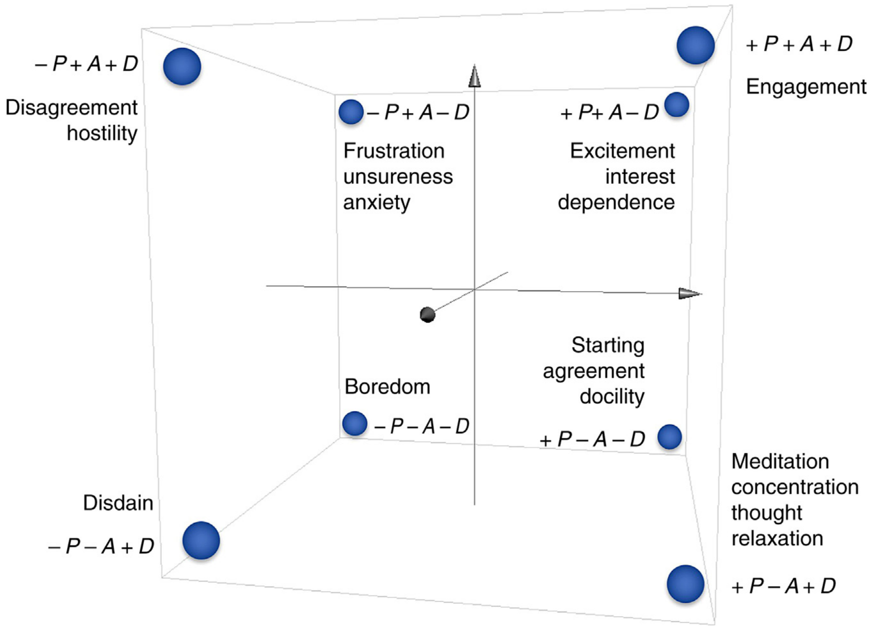

FIGURE 11.2 **Pleasure-Arousal-Dominance dimensional model.** *Source: From Gonzalez Sanchez, J., 2016. Affect-Driven Self-Adaptation: A Manufacturing Vision with a Software Product Line Paradigm. Doctoral dissertation. Arizona State University, Arizona, United States.*

Mehrabian (1996) proposed expanding the two-dimensional model to three dimensions by adding another axis: dominance. Dominance represents how controlling and dominant versus how controlled or submissive one feels. For instance, while both frustration and disagreement are unpleasant emotions, disagreement is a dominant emotion and frustration is submissive. Fig. 11.2 shows the three-dimensional model and plots some emotions with their pleasure–arousal–dominance (PAD) vectors.

Since continuous values characterize the measure of affect during human interactions, Mehrabian's model, also known as the PAD model, has been recommended for real-time emotion recognition. Gilroy et al. (2009) describe a case study using the PAD model in real-time for an art installation.

GATHERING DATA: APPROACHES AND TECHNOLOGIES

Having reviewed the conceptual background to affect measurement, our roadmap starts by describing some popular, inexpensive, easy to install, and widely available sensing approaches and technologies that we have been using in our research over the last 4 years. For each of them, we describe the data gathered and its characteristics, as well as outlining its pros and cons.

Brain–Computer Interfaces

BCI is a physiological instrument that uses brainwaves as data sources. Most BCIs work under the principles of electroencephalography (EEG), recording electrical activity along the scalp produced by the firing of neurons within the brain. BCI devices of varying accuracy and cost are widely available. Three devices that we have had the opportunity to work with are described below:

1. *NeuroSky biosensor*. This device facilitates low-cost EEG-linked research using one dry sensor, situated at the forehead (Fig. 11.3A), which provides a very easy and almost nonintrusive setup. It provides raw data at a sampling rate of 512 Hz. Its software is able to extract constructs for attention, meditation, and eye blinking, as well as delta, theta, low alpha, high alpha, low beta, high beta, and gamma waves.
2. *Emotiv EPOC headset*. Emotiv raw data output includes 14 values, 7 channels for each brain hemisphere. Electrodes for these channels are situated and labeled according to the CMS/DRL configuration: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4 (American Electroencephalographic Society, 1990), as shown in Fig. 11.3B. Additionally, two accelerometers track head movements; they are reported as AccX and AccY. Accelerometer values can be used to identify nodding, headshaking, or yes/no indication. Its setup is straightforward and can be accomplished in as little as 2–3 min with one caveat: users with thick or curly hair may present a challenge. The device requires that the electrodes be coated in a multipurpose solution. They normally remain wet for about 1 h before needing to be remoistened. While using this headset, it is important that

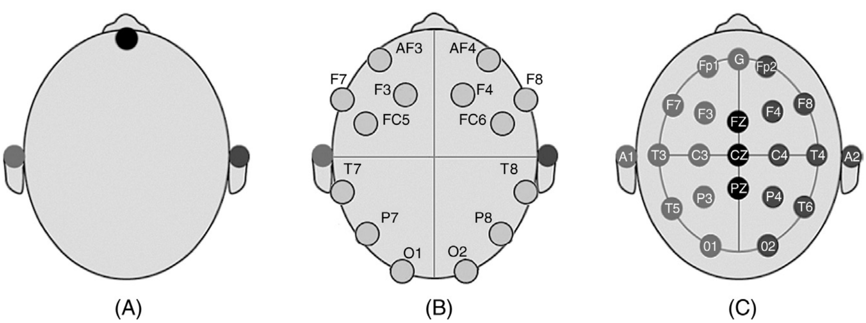

FIGURE 11.3 BCI devices range from simple to complex; from (A) 1 channel at the forehead, NeuroSky; to (B) 14 CMS/DRL channels, Emotiv EPOC; to (C) standard international 10-20 configuration, ABM B-Alert X-Series.

the user does not chew gum since chewing generates noise signals. The device reports raw data at a sampling rate of 128 Hz, reporting packets of 128 samples each time. This system is able to infer five emotional constructs (engagement, boredom, excitement, frustration, and meditation) and to detect a wide range of facial gestures (blink, wink left and right, look left and right, raise brow, furrow brow, smile, clench, smirk left and right, and laugh). Emotional constructs and facial gestures are reported at a sampling rate of 8 Hz.

3. *ABM B-Alert X-Series EEG systems*. B-Alert X-Series systems[a] are appropriate for the high-quality, nonmedical wireless acquisition of EEG and physiological signals. It generates validated cognitive state metric and cognitive workload metric computations in real-time or during offline analyses. It applies sensors according to the standard international 10–20 system, as shown in Fig. 11.3C. It reports data at a sampling rate of 256 Hz. Its setup requires the application of conductive foam between the electrodes and the scalp. Similar to the Emotiv system, thick or curly hair can represent a challenge to fitting the electrodes. The setup of the headset includes an impedance test (5 min) to test each node's connection with the scalp, as well as a baseline test (15–20 min) to normalize the system for each individual user. The latter comprises a vigilance task, in which the user responds to visual stimuli, and an audio task, in which the user reacts to audio tones. A user's baseline test results can be saved and reused. Due to the sensitivity of this headset, the B-Alert User Manual (version 2.0 from 2011) recommends ensuring that the user is well rested (not sleepy) and that the user did not consume nicotine or caffeine immediately prior to the experiment.

Table 11.1 contains an example of raw data collected from a BCI, specifically an extract of a dataset collected using the Emotiv EPOC headset. Note that a timestamp is included for each row. The timestamp is a 15-digit string in which each consecutive pair of numbers respectively indicates the year, month, day, hour, minute, and second value, and the final three digits indicate the millisecond value. For instance, the timestamp in the first row in Table 11.1 is 141116112544901, which denotes November 16, 2014 at 11:25:44.901 a.m. This format is followed in all data reports presented in this chapter. The Emotiv EPOC headset reports raw data in packets of 128 samples and all samples in a packet are labeled with the same timestamp. Thus, for this device, a change in the timestamp occurs every 128 rows.

Table 11.2 shows a sample dataset of the affective constructs produced by Emotiv software. Notice that the construct values are expressed as probabilities (values ranging from 0 to 1). The construct excitement is reported

[a]http://www.biopac.com/B-Alert-X10-Analysis-Software

TABLE 11.1 Extract of Raw Data Collected Using the Emotiv EPOC Headset

Timestamp	AF3	F7	F3	FC5	T7	P7	O1	O2	P8	T8	FC6	F4	F8	AF4	AccX	AccY
141116112544901	4542.05	4831.79	4247.18	4690.26	4282.56	4395.38	4591.79	4569.23	4360	4570.77	4297.44	4311.28	4282.56	4367.18	1660	2003
141116112544901	4536.92	4802.05	4243.08	4673.85	4272.31	4393.33	4592.82	4570.26	4354.87	4570.26	4292.31	4309.74	4277.95	4370.77	1658	2002
141116112545010	4533.33	4798.97	4234.87	4669.74	4301.03	4396.92	4592.31	4570.77	4351.28	4561.03	4281.54	4301.54	4271.28	4363.59	1659	2003
141116112545010	4549.23	4839.49	4241.03	4691.28	4333.85	4397.95	4596.41	4567.18	4355.9	4556.41	4286.15	4306.15	4277.95	4369.74	1659	2003
141116112545010	4580	4865.64	4251.79	4710.26	4340	4401.54	4603.59	4572.82	4360	4558.46	4298.97	4324.62	4296.41	4395.9	1657	2004
141116112545010	4597.44	4860	4252.82	4705.64	4350.26	4412.31	4603.59	4577.44	4357.44	4555.9	4295.38	4329.23	4296.41	4414.36	1656	2005
141116112545010	4584.62	4847.69	4246.67	4690.26	4360	4409.23	4597.44	4569.74	4351.79	4549.74	4278.97	4316.92	4272.82	4399.49	1656	2006
141116112545010	4566.15	4842.05	4238.46	4684.1	4322.05	4389.74	4592.82	4566.67	4351.79	4549.74	4274.36	4310.26	4262.05	4370.77	1655	2005
141116112545010	4563.59	4844.62	4231.79	4687.69	4267.69	4387.69	4594.36	4580	4361.03	4556.41	4278.97	4310.77	4274.36	4370.77	1653	2006
141116112545010	4567.18	4847.18	4233.33	4688.72	4285.13	4409.23	4602.05	4589.23	4368.21	4560	4280.51	4310.77	4281.54	4390.26	1655	2004

TABLE 11.2 Extract of Affective Constructs Reported by the Emotiv EPOC Headset

Timestamp	Short-term excitement	Long-term excitement	Engagement	Meditation	Frustration
141116091145065	0.447595	0.54871	0.834476	0.333844	0.536197
141116091145190	0.447595	0.54871	0.834476	0.333844	0.536197
141116091145315	0.447595	0.54871	0.834476	0.333844	0.536197
141116091145440	0.487864	0.546877	0.834146	0.339548	0.54851
141116091145565	0.487864	0.546877	0.834146	0.339548	0.54851
141116091145690	0.487864	0.546877	0.834146	0.339548	0.54851
141116091145815	0.487864	0.546877	0.834146	0.339548	0.54851
141116091145940	0.521663	0.545609	0.839321	0.348321	0.558228
141116091146065	0.521663	0.545609	0.839321	0.348321	0.558228
141116091146190	0.521663	0.545609	0.839321	0.348321	0.558228

twice: once as short-term excitement, reflecting an immediate change, and then as long-term excitement, using the cumulative readings over time to calculate a value that reflects the overall change. Timestamps confirm a sampling rate for affective constructs at 8 Hz (one sample every 125 ms).

In general, BCIs are easy to use and portable. However, it is worth noting that they require time for setup and calibration (from an average of 10 min for the Emotiv headset to an average of 40 min for the ABM B-Alert), and variables, such as the battery level of the headset, the noise in the environment, and the connection between the electrodes and the scalp should be monitored continuously. Furthermore, the headset can only be a limited distance from the host computer, and the number of headsets that can operate in the same room without creating interference is also a limited. The presence of other wireless devices using the same bandwidth can also produce interference. Ongoing research includes the development of EEG devices that use dry electrodes; these are becoming more readily available and overcome the limitations of scalp preparation and wet gels. Cognionics and Wearable Sensing are examples of companies working on improving this aspect of this device.

Facial Gestures

Emotion recognition systems based on facial gesture enable real-time analysis, tagging, and inference of cognitive affective states from a video recording of the face. It is assumed that facial expressions are triggered for a period of time when an emotion is experienced and so emotion detection can be achieved by detecting the facial expression related to it. From each

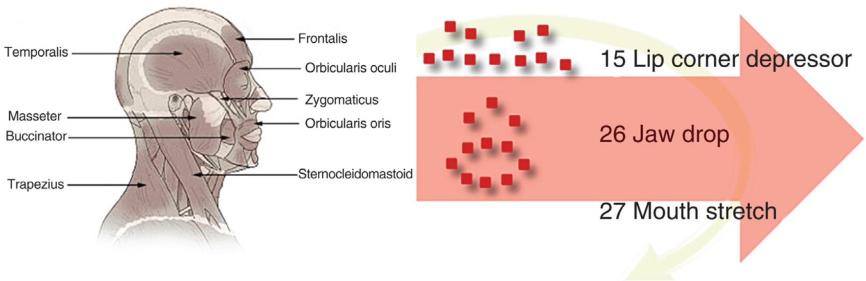

FIGURE 11.4 **Movements of individual facial muscles are encoded as action units. For instance, lip corner depression is coded as action unit 15, jaw drop as action unit 26, and mouth stretch as action unit 27.**

facial expression, a set of facial action units is extracted, each facial action unit identifying an independent motion of the face. Movements in facial muscles are perceived as changes in the position of the eyes, nose, and mouth. Computer systems implement this approach by capturing images of the user's facial expressions and head movements. Those systems detect changes in the position of the eyes, nose, and mouth as changes in the position of dots in a coordinate system. Then, by analyzing those changes, the occurrence of a facial action unit can be determined. The Facial Action Coding System (FACS) documents 46 possible facial action units (Ekman et al., 1980). For instance, happiness is associated with the occurrence of action units 6 and 12 (cheek raiser and lip corner puller), and sadness is associated with the occurrence of action units 1, 4, and 15 (inner brow raiser, brow lowerer, and lip corner depressor). Fig. 11.4 schematizes the process.

Diverse methodologies are used to infer affect from facial images, some of which are summarized in Table 1 of Calvo and D'Mello (2010). Examples of emotion recognition systems based on facial gesture that conduct real-time frame-by-frame facial expression recognition from a video stream include:

1. *MindReader* (Kaliouby and Robinson, 2005) uses a standard 30 fps USB webcam to capture the user's face. It also takes into account an analysis of head and shoulder movements. It provides results at a sampling rate of 10 Hz and is able to infer the affective states of agreement, concentration, disagreement, interest, thought, and unsureness.
2. *iMotion FACET*[b] uses a standard 30 fps USB webcam to capture the user's face. It provides results at a sampling rate of 30 Hz and is able to infer nine states: seven basic emotions (joy, anger, fear, sadness, disgust, surprise, and contempt) and two complex emotions (confusion and frustration).

Table 11.3 shows a sample dataset for affect constructs provided, in real-time, by MindReader software. The constructs are probabilities (values

[b]http://imotionsglobal.com

TABLE 11.3 Extract of Affective Constructs Reported by MindReader

Timestamp	Agreement	Concentration	Disagreement	Interest	Thought	Unsureness
141116112838516	0.001836032	0.999917000	0.000179000	0.164854060	0.571142550	0.045950620
141116112838578	0.001447654	0.999951600	0.000129000	0.163106830	0.595892100	0.042706452
141116112838672	0.000597000	0	0.000150000	0.449962940	0.455276130	0.007896970
141116112838766	0.000246000	0	0.000175000	0.774456860	0.321447520	0.001418217
141116112838860	0.000101000	0	0.000204000	0.935119150	0.211671380	0.000253000
141116112838953	0.000041800	0	0.000238000	0.983739000	0.132086770	0.000045200
141116112839016	0.000017200	0	0.000278000	0.996077400	0.079410380	0.000008070
141116112839110	0.000007100	0	0.000324000	0.999062660	0.046613157	0.000001440
141116112839156	0.000002920	0	0.000377000	0.999776540	0.026964737	0.000000257
141116112839250	0.000001210	0	0.000440000	0.999946700	0.015464196	0.000000046

ranging from 0 to 1). The timestamp column shows an approximate sampling rate of 10 Hz (one sample every 100 ms).

The use of emotion recognition systems based on facial gesture has a number of advantages. They are largely nonintrusive because they do not involve attaching sensors to the user and are inexpensive because they do not require expensive hardware. They also have reasonable accuracy, although the accuracy is challenged when the user's movements cause the face to turn away from the camera's line of sight, such as when a user lies back in the chair or looks down to read or write. Cohn and De la Torre (2014) describe in more detail the challenges associated with automated facial image analysis.

Eye Tracking

Eye-tracking systems measure eye position, eye movement, and pupil size to detect zones in which the user has a particular interest at a specific time. There are a number of methods for measuring eye movement. The most popular are optical methods, in which light, typically infrared, is reflected from the eye and sensed by a camera or some other specially designed optical sensor. The data is then analyzed to extract eye rotation from changes in the reflections. Optical methods are widely used for gaze tracking and are favored for being noninvasive and inexpensive. An example of a commercial optical eye-tracking system is the Tobii T60XL Eye Tracker. The Tobii Eye Tracker reports data at a sampling rate of 60 Hz, and the reported data includes attention direction as a gaze point (x and y coordinates), duration of fixation, and pupil dilation.

Pupil diameter has been demonstrated to be an indicator of emotional arousal, as seen in Bradley et al. (2008), who found that pupillary changes were larger when viewing emotionally arousing pictures, regardless of whether these were pleasant or unpleasant. Pupillary changes during picture viewing covaried with skin conductance changes, supporting the interpretation that sympathetic nervous system activity modulates these changes.

A sample dataset from a Tobii T60XL Eye Tracker is shown in Table 11.4. Gaze point values (GPX and GPY columns) range from 0 to the size of the display; pupil (left and right) is the size of the pupil in millimeters; validity (left and right) is an integer value ranging from 0 to 4 (0 if the eye is found and the tracking quality is good and 4 if the eye cannot be located by the eye tracker); and fixation zone is a sequential number corresponding to one or a set of predefined zones in which special interest exists. Timestamps in the table confirm a sampling rate of 60 Hz (approximately one sample every 16 ms).

Eye-tracking systems can be fixed (embedded in a display), mobile (able to be connected and mounted in diverse displays), or wearable

TABLE 11.4 Extract of Data Collected Using Tobii T60XL Eye Tracker

Timestamp	GPX	GPY	Pupil left	Validity *L*	Pupil right	Validity *R*	Fixation zone
141124162405582	636	199	2.759313	0	2.88406	0	48
141124162405599	641	207	2.684893	0	2.855817	0	48
141124162405615	659	211	2.624458	0	2.903861	0	48
141124162405632	644	201	2.636186	0	2.916132	0	48
141124162405649	644	213	2.690685	0	2.831013	0	48
141124162405666	628	194	2.651784	0	2.869714	0	48
141124162405682	614	177	2.829281	0	2.899828	0	48
141124162405699	701	249	2.780344	0	2.907665	0	49
141124162405716	906	341	2.853761	0	2.916398	0	49
141124162405732	947	398	2.829427	0	2.889944	0	49

(embedded in a pair of glasses). Regardless of the type of system, the setup process is fairly easy. The calibration process includes having the user follow an object around the display area with their eyes (for embedded and mobile systems), or having them stare at a particular point (for wearable glasses). The calibration for embedded and mobile systems requires time to ensure that the eyes of the user are within the line of sight of the infrared and optical sensors and that nothing is producing glare for the camera, which could affect the reflection and thus the tracking of eye movements. The reliability of embedded and mobile systems is reduced by glare on the cameras, the incorrect position of the user's face, and the presence of framed glasses or eye disorders, such as strabismus. In the case of systems in glasses, important things to consider are the interference of other wireless devices and the distance that the glasses can be from the host computer.

Physiological Sensor: Skin Conductance

Electrodermal activity (EDA) is the continuous variation in the electrical characteristics of the skin, which varies with the moisture level. The moisture level depends on the sweat glands and blood flow, which are controlled by the sympathetic and parasympathetic nervous systems. Although the electrical variation alone does not identify a specific emotion, a relationship has been established between this and emotional arousal, that is, how energized or soporific the user feels. A skin conductance device senses EDA by measuring the conductance of the skin. To measure skin conductance, the electrical resistance between two electrodes, normally attached to the skin about an inch apart, when a very weak current is steadily passed between them is measured. Skin conductance is perhaps the most inexpensive method discussed here in terms of the hardware and software required. Examples of skin conductance sensors used in our research are:

1. A wireless Bluetooth device that reports data at a sampling rate of 2 Hz, designed by MIT Media Lab (Strauss et al., 2005) and modified in-house at Arizona State University. The data reported includes the battery level, a floating-point value ranging from 0 to 5, and the conductance, a floating-point value with a lower value of 0, which increases in proportion to the increase in the skin conductance.
2. The Shimmer3 GSR+[c] monitors skin conductance using two reusable electrodes attached to two fingers (middle and ring). It is able to report data at a variable sampling rate in the range of hundreds of hertz. Additionally, the Shimmer3 GSR+ provides an estimated heart

[c]http://www.shimmersensing.com

TABLE 11.5 Extract of Data Collected Using an in-House Built Skin Conductance Device

Timestamp	Voltage	Conductance
141116101332262	2.482352941	1.030696176
141116101332762	2.482352941	1.023404165
141116101333262	2.482352941	1.019813274
141116101333762	2.482352941	1.041657802
141116101334247	2.482352941	0.998280273
141116101334747	2.482352941	0.991181142
141116101335247	2.482352941	0.980592229
141116101335747	2.482352941	0.998280273
141116101336247	2.482352941	1.012586294
141116101336762	2.482352941	1.012586294

rate calculated from a photoplethysmogram signal captured by a wired ear or finger clip electrode.

A sample dataset from our in-house developed skin conductance device is shown in Table 11.5. The timestamp confirms a precise sampling rate of 2 Hz. Feature selection and extraction approaches for this measurement can be consulted in Strauss et al. (2005) and Cooper et al. (2009).

The setup for a skin conductance device does not require calibration and as long as its battery has sufficient power, the data will be gathered accurately.

Physiological Sensor: Heart Rate Variability

Heart rate is defined as the number of heartbeats occurring per minute and the average resting heart rate for an adult human is between 60 and 90 beats. The heart rate goes up when activity in the sympathetic nervous system (which controls the body's involuntary responses to a perceived threat) increases. The heart rate goes down when activity in the sympathetic nervous system decreases. Conversely, the heart rate goes up when parasympathetic nervous system activity (which controls the body's involuntary responses at rest) decreases (because there is less inhibition). The heart rate goes down when parasympathetic nervous system activity increases (because there is more inhibition). Although HRV is influenced by numerous physiological and environmental factors, the influence of the autonomic nervous system on cardiac activity is particularly prominent and of psychophysiological importance. HRV analysis is emerging as an objective measure of a regulated emotional response. HRV measures the variation in the time intervals between heartbeats and has been related to emotional arousal. Experimentation and theory support the usefulness of HRV as an objective index of the brain's ability to organize regulated emotional responses and as a marker of individual differences in emotion regulation capacity (Appelhans and Luecken, 2006).

One recommendation for this device's use is that the sensor should be placed on the nondominant hand or ankle to minimize noise caused by movement. We are currently introducing this method to our toolkit using a Shimmer3 GSR+ sensor and looking forward to a comparative analysis of the results against other sensing systems.

Body Language: Pressure

Like facial gestures, body language is a potential channel of emotional expression. Body language can be more challenging to identify than facial gestures. However, several approaches have been developed to overcome this challenge. One uses pressure sensors. Pressure sensors allow the amount of pressure applied to an object, such as a mouse, a game controller, or a keyboard, to be detected. Pressure has been correlated to levels of frustration (Qi and Picard, 2002). Qi and Picard created a device in which pressure sensors reporting data at a sampling rate of approximately 6 Hz were embedded in a mouse. The six sensors were situated in the right, left, and middle front and rear parts of the mouse. The raw data from the six pressure sensors was processed and correlated to levels of frustration (Cooper et al., 2009). The raw pressure data is represented by integer values in the range of 0 to 1024, where 0 represents the highest pressure. The machine learning approach for inferring affect based on measurements from such pressure values can be consulted in Qi and Picard (2002) and Cooper et al. (2009).

A sample dataset from an in-house version of Qi and Picard's device with six pressure sensors is shown in Table 11.6. The timestamp shows a sampling rate of approximately 6 Hz.

Pressure sensors are nonintrusive since the user does not have to wear any equipment and the device does not require any calibration. In addition, the functionality of the target object is not affected by the introduction of the pressure sensors, although the pressure sensors in the object must be connected to the computer using a serial or USB port.

Body Language: Posture

Another approach to body language analysis is posture detection. One approach that we use to achieve this consists of using a low-cost, low-resolution pressure sensitive seat cushion and back pad. Sensors are embedded in the seat cushion and back pad so that they are imperceptible

TABLE 11.6 Extract of Data Collected Using a Mouse Enhanced With Pressure Sensors

Timestamp	Right rear	Right front	Left rear	Left front	Middle rear	Middle front
140720113306312	1023	1023	1023	1023	1023	1023
140720113306468	1023	1023	1023	1023	1023	1023
140720113306625	1023	998	1023	1002	1023	1023
140720113306781	1023	1009	1023	977	1023	1023
140720113306937	1023	794	1023	982	1023	1023
140720113307109	1023	492	1022	891	1023	1023
140720113307265	1023	395	1021	916	1019	1023
140720113307421	1023	382	1021	949	1023	1023
140720113307578	1023	364	1022	983	1023	1023
140720113307734	1023	112	1021	1004	1023	1023

to the user. They are positioned in a triangle configuration in the middle, right, and left. An example of this device was developed at Arizona State University based on the design of a more expensive high-resolution unit from the MIT Media Lab (Mota and Picard, 2003). The data from three pressure sensors in the back, three in the seat, and the two accelerometers is processed to obtain the net seat change, net back change, and sit forward features. Those features are then used to infer a level of interest. The machine learning approach for inferring affect related measures from such values is described in Mota and Picard (2003) and Cooper et al. (2009).

A sample dataset from a posture device is shown in Table 11.7. As is the case with the pressure sensor, the values range from 0 to 1024; however, in this instance 1024 represents the highest pressure.

Posture sensing using a seat cushion enhanced with pressure sensors is an inexpensive, easy to implement, and nonintrusive measurement, although, like the pressure sensor, the seat cushion must be connected to the computer using a serial or USB port. Moreover, these devices do not require any calibration.

Language Processing: Writing Patterns

Text-based human interaction often carries important emotional significance. In the context of computer-mediated communication, some methods for text-based emotion recognition have been developed, such as the one described in Krcadinac et al. (2013). They work at the sentence level

TABLE 11.7 Extract of Data Collected Using a Seat Cushion Enhanced With Pressure Sensors

Timestamp	Right seat	Middle seat	Left seat	Right back	Middle back	Left back	Net seat change	Net back change	Sit forward
140720074358901	1015	1019	1012	976	554	309	12	152	0
140720074359136	1008	1004	1012	978	540	305	22	20	0
140720074359401	1015	1012	1008	974	554	368	19	81	0
140720074359636	1001	1004	1016	975	548	306	30	69	0
140720074359854	1015	1011	1003	967	559	418	34	131	0
140720074400120	1011	1008	1001	968	620	358	9	122	0
140720074400354	1011	1011	1013	968	541	413	15	134	0
140720074400589	1012	1010	1006	974	565	314	9	129	0
140720074400839	1016	1014	1012	972	668	290	14	129	0
140720074401089	1012	1012	1004	858	108	2	14	962	0

and implement a recognition technique founded on a refined keyword spotting method, which employs a set of heuristic rules, a WordNet-based word lexicon, and a lexicon of emoticons and common abbreviations. Their approach is implemented through a software system named Synesketch, released as a free, open source software library. Synesketch analyses the emotional content of text sentences in terms of emotional types (happiness, sadness, anger, fear, disgust, and surprise), weights (how intense the emotion is), and valence (positive or negative).

Other approaches go beyond word- or sentence-level analysis and perform a semantic analysis of the text. These approaches move toward sentiment analysis and opinion mining, which use natural language processing, text analysis, and computational linguistics to identify and extract subjective information from source materials. They aim to determine the attitude of a writer with respect to some topic or the overall contextual polarity of a document. The attitude may be the writer's emotional state, or the intended effect the writer wishes to have on the reader. A summary of diverse approaches to inferring affect from text analysis are summarized in Table 4 in Calvo and D'Mello (2010).

No calibration is required for emotion recognition from writing patterns since it does not involve any particular hardware or device; this is fully implemented through software. To our knowledge, currently available options are limited to the English language. Even though Synesketch is open source and free, commercial implementation may be expensive as it is comparable in cost to facial recognition systems.

DATA HANDLING: SAMPLING, FILTERING, AND INTEGRATION

Section "Gathering Data: Approaches and Technologies" presents a selection of the broad range of technologies and approaches for sensing and gathering data correlated with affective state changes, which are becoming increasingly more accessible and robust. Our roadmap now turns to explore the problems and methodologies associated with sampling, filtering, and integrating affective data.

Sampling

The first challenge when dealing with multiple sources of data relates to the sampling rate. Sampling rates for different sources vary from a few samples per second (2 Hz for skin conductance sensing) to over a hundred samples per second (128 Hz for raw brain wave data). Thus, a decision must be made regarding how to synchronize and integrate the data; common options include:

1. sampling at the lowest rate;
2. sampling at the highest rate; and
3. determining a suitable sampling rate between the computed lower and higher limits, prioritizing specific trade-offs.

Using the lowest rate means losing data and potentially forfeiting accuracy; however, it involves less computer strain, more latency, and using less hard drive space. While using higher rates potentially provides greater accuracy, it does involve more computer strain, less latency, and more hard drive space. The higher sampling rate is more appropriate if the goal is to create a model for off-line analysis. For real-time analysis, it may be preferable to determine a suitable sampling rate according to the system requirements and the computational resources. In real-time systems emotions experienced by the user for substantial periods of time are usually the priority, rather than emotions that occur briefly.

Filtering

One of the most relevant goals of filtering data is noise reduction. Noisy data are unwanted readings that might impact the efficiency of the data processing. Ways in which sensors can report noisy data include:

1. MindReader facial recognition software reports a value of −1 when an affective state inference cannot be achieved; this usually occurs when the face moves out of the camera's line of sight.
2. Eye tracker reports gaze point values as negative or greater than the screen resolution values when light conditions are bad or the user's

head movements interfere with the reading. The validity attribute indicates how reliable an eye tracker sample is. Usually, samples with a validity value of 0 to 3 (high and medium quality) are retained and samples with a validity value of 4 are discarded.

3. BCIs, such as the Emotiv headset, report low-quality data when the electrodes become dry and when nearby electronic devices affect the EEG signal. Since a baseline of 4000 is defined, noise values can be detected and excluded.
4. For skin conductance devices, the battery level is important. Low battery power reduces the quality of the conductance readings.

Various techniques are employed to reduce noise by cleaning unwanted values from the data. Common filtering approaches for denoising data are further discussed by Manolakis et al. (2000). These techniques include:

1. Low-pass filtering. This ignores signals with a frequency lower than a certain cutoff value and attenuates signals with frequencies higher than the cutoff value to understand the main signal behavior.
2. Moving average filtering. This analyzes data points by creating a series of averages for different subsets of the full data set.
3. Median filtering. This denoises data in a similar way to moving average filtering but uses the median value rather than an average. In some cases, these results are more robust because the median value is less likely to be affected by outlying (unusual) values.

Integration

Integration is required when a multimodal approach is implemented since it takes into account the diverse sources of data that contribute to the inference process. Since each source may capture data of a different type at a different sampling rate, it is a challenge to combine the data to create a single enhanced dataset combining all records into one, with a column for each feature and a row for each timestamp. A row contains the values recorded for any source that corresponds with that timestamp. Examples of integration methods include sparse data, state machine, interpolation, and mapping to a coordinate model, each of which is presented briefly below:

Sparse data. If no value exists from a given source for a given timestamp, the cell remains empty. For instance, in the first row of Table 11.8, for the timestamp 141014135755652, there was no data from the eye tracker but there was data from the Emotiv EPOC headset and the skin conductance sensor. A similar situation occurred in the second row, where only eye tracker data exists for the timestamp 141014135755659. The presence of numerous empty cells in the dataset is referred to as sparse data. Applying a high sampling rate usually results in sparse data.

TABLE 11.8 Integration of Measurements (From an Eye Tracker, an Emotiv EPOC Headset, and a Skin Conductance Device) Using a Sparse Data Approach

Timestamp	Fixation	GPX	GPY	Short-term excitement	Long-term excitement	Engagement/ boredom	Meditation	Frustration	Conductance
141014135755652				0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755659	213	573	408						
141014135755668				0.436697	0.521059	0.550011	0.335825	0.498908	
141014135755676	213	566	412						
141014135755692	213	565	404						
141014135755709	213	567	404						
141014135755714									
141014135755726	213	568	411						
141014135755742	213	568	409						
141014135755759	213	563	411						
141014135755761									
141014135755776	213	574	413						
141014135755792	213	554	402						
141014135755809	214	603	409						
141014135755824									
141014135755826	214	701	407						
141014135755842	214	697	403						

141014135755859	214	693	401						
141014135755876	214	700	402						
141014135755892	214	701	411						
141014135755909	214	686	398						
141014135755918									
141014135755926	214	694	399						
141014135755942	214	694	407						
141014135755959	214	698	404						
141014135755964									
141014135755976	214	704	398						
141014135755992	214	693	415						
141014135756009	214	696	411						
141014135756025	215	728	406						
141014135756027				0.436697	0.521059	0.150011	0.335825	0.998908	

State machine. Instead of considering an affective state to exist only in the period of time equivalent to the sample, this method assumes that the state persists until a new state is obtained; thus, a source is in a state until it changes from that state to another. An example is shown in Table 11.9. There it is evident that, due to the sampling rate difference between the eye tracker (60 Hz), the Emotiv system (8 Hz), and the skin conductance device (2 Hz) in the interval between 57:55.652 and 57:56.027, there are 23 values reported from the eye tracker, 3 from the Emotiv system, and only 1 from the skin conductance device. Using the state machine approach, the values are repeated for each time stamp until a new value is obtained and there are no empty cells. The state machine approach is a good choice for real-time integration analysis.

Interpolation. This is a method for constructing new data points within the range of a discrete set of known data points. It is commonly used to deal with missing values by predicting them. Missing value computation is an active research topic and there are several successful techniques for handling the missing values in the literature (Little and Rubin, 2002).

Mapping to a coordinate model. This works with affective constructs and consists of mapping constructs to a coordinate vector and then adding all the coordinate vectors together. This approach has been suggested for systems that perform emotion recognition in real-time and use their inferences to trigger adaptive changes in a target system. Gilroy et al. (2009) suggested the use of the PAD three-dimensional coordinate model. The process of mapping the data from row 2 in Table 11.9 to the PAD three-dimensional coordinate model is described later. To simplify our example, only constructs from the Emotiv EPOC headset and the skin conductance device are used, that is, data from the eye tracker is ignored. Also, assuming a focus on real-time systems, short-term excitement is used here and long-term excitement is ignored.

Each construct value is mapped to a PAD coordinate vector as follows:

1. Excitement corresponds to $[+P, +A, -D]$, thus a value of 0.436697 can be mapped to the coordinate vector [0.436697, 0.436697, −0.436697].
2. Engagement and boredom are reported in the same column. Engagement corresponds to $[+P, +A, +D]$ and boredom corresponds to $[-P, -A, -D]$. A value of 0.5 means equilibrium, values below 0.5 are indicators of boredom, and values above 0.5 are indicators of engagement; thus, engagement/boredom can be mapped, in a scale of −1 to 1, to [(value − 0.5)*2, (value − 0.5)*2, (value − 0.5)*2]. Then a value of 0.550011 is mapped to [0.100022, 0.100022, 0.100022], which is positive and thus corresponds to a low level of engagement.
3. Meditation corresponds to $[+P, -A, +D]$; thus, a value of 0.335825 can be mapped to the vector [0.335825, −0.335825, 0.335825].

TABLE 11.9 Integration of Measurements (From an Eye Tracker, an Emotiv EPOC Headset, and a Skin Conductance Device) Using a State Machine Approach

Timestamp	Fixation	GPX	GPY	Short-term excitement	Long-term excitement	Engagement/ boredom	Meditation	Frustration	Conductance
141014135755652	213	574	414	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755659	213	573	408	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755668	213	573	408	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755676	213	566	412	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755692	213	565	404	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755709	213	567	404	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755714	213	567	404	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755726	213	568	411	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755742	213	568	409	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755759	213	563	411	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755761	213	563	411	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755776	213	574	413	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755792	213	554	402	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755809	214	603	409	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755824	214	603	409	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755826	214	701	407	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628

(Continued)

TABLE 11.9 Integration of Measurements (From an Eye Tracker, an Emotiv EPOC Headset, and a Skin Conductance Device) Using a State Machine Approach *(cont.)*

Timestamp	Fixation	GPX	GPY	Short-term excitement	Long-term excitement	Engagement/ boredom	Meditation	Frustration	Conductance
141014135755842	214	697	403	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755859	214	693	401	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755876	214	700	402	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755892	214	701	411	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755909	214	686	398	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755918	214	686	398	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755926	214	694	399	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755942	214	694	407	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755959	214	698	404	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755964	214	698	404	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755976	214	704	398	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135755992	214	693	415	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135756009	214	696	411	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135756025	215	728	406	0.436697	0.521059	0.550011	0.335825	0.498908	0.401690628
141014135756027	215	728	406	0.436697	0.521059	0.150011	0.335825	0.998908	0.401690628

4. Frustration corresponds to [−*P*, +*A*, −*D*]; thus, a value of 0.498908 can be mapped to the vector [−0.498908, 0.498908, −0.498908].
5. Skin conductance is not mapped to an affective construct but to arousal; thus, a normalized value of 0.401690628 can be mapped as zero pleasure, positive arousal, and zero dominance, that is, to the vector [0, 0.401690628, 0].

Adding the vectors gives a resultant vector of [0.373636, 1.101492628, −0.499758], a vector in the zone of [+*P*, +*A*, −*D*], that is, a level of excitement. This vector represents the cumulative affective state at timestamp 141014135755659 (Oct 14, 2014 13:57:55.659).

DATA ANALYSIS

Once time sampling has been accounted for, noise has been removed, missing data values have been addressed, and the data has been integrated, data analysis can start. There is a vast body of literature available on the topic of data analysis. This section describes the overall process, and exemplifies techniques for data analysis in general and for the correlation of affect measurements with stimuli. Two examples of correlating affect and stimuli off-line using data collected in experimental studies are presented using two free tools. Advanced methods for data analysis are surveyed in Novak et al. (2012).

The overall strategy for data analysis is depicted in Fig. 11.5 and can be summarized as follows:

1. *Feature extraction*. This step starts with a set of clean data and builds derived values (called features) intended to be informative and

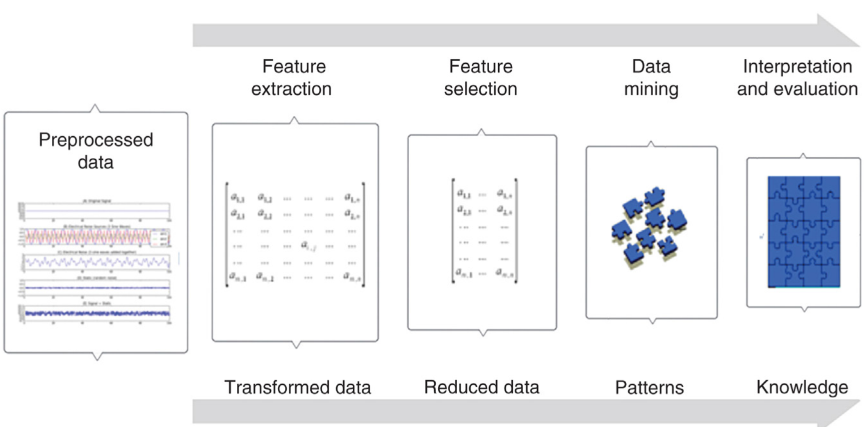

FIGURE 11.5 **Steps for data analysis: feature extraction, feature selection, data mining, and interpretation and evaluation.**

nonredundant. The extracted features are expected to contain the relevant information from the input data, so that subsequent steps can be performed using this reduced representation instead of the entire initial set of data. With dimensionality reduction or transformation methods, the effective number of variables under consideration can be reduced, or invariant representations for the data can be found.

2. *Feature selection*. This step entails the process of selecting a subset of relevant features for use in the construction of a model. The central premise of using a feature selection technique is that the data contains many features that are either redundant or irrelevant, and can thus be removed without incurring much loss of information. Feature selection techniques allow: (1) the simplification of models for ease of interpretation, (2) shorter training times, and (3) enhanced generalization by reducing overfitting.
3. *Data mining*. This step involves the selection of a mining approach to identify patterns of interest. Common approaches include classification rules or trees, regression, and clustering. The selection process takes into account: (1) the parameters that are going to be used and which model might be appropriate for them, for instance, models for categorical data are different to models for vectors; and (2) the overall criteria of the knowledge discovery process, for instance, understanding the model or achieving outstanding predictive capabilities.
4. *Interpretation and evaluation*. Interpretation involves the visualization of the extracted patterns and models, or the visualization of the data given the extracted models. Evaluation is the application of the knowledge by incorporating it into another system for further actions, or simply documenting it and reporting it to interested parties.

The following two subsections outline and exemplify two approaches for data analysis, regression and classification, using two freely available tools, *Eureqa* and *Weka*, respectively.

Regression Example

For regression or reverse engineering data searches, we use *Eureqa*[d]. *Eureqa* is used to compute mathematical expressions for structural relationships in the data records (Dubcakova, 2011). Typically, the records hold information about the actions or behaviors and affective states of a user who was engaged in an experimental setting. For instance, consider exploring the relationships between affective states and the tendency to make mistakes. We collected data from participants engaged in playing the video game Guitar Hero. The goal of the game is to press one or more colored buttons at the same time as moving target lights of the same color cross a line

[d]http://www.nutonian.com/products/eureqa/

on the screen. In each 1-h session, the first 15 min were allocated to practice, followed by a 45-min session in which participants played four songs of their choice, one of each level: easy, medium, hard, and expert. The data collected included the errors made and constructs for engagement, excitement, meditation, and frustration. For this example, a partial dataset was fed to the *Eureqa* tool and used to create a model expressed as a mathematical equation, showing possible relationships between attributes, where sensor readings are the independent variables and errors, the number of mistakes that are likely to happen in a timeframe, are the target variable.

Eureqa automatically splits the dataset in two for training and validation, respectively. The training data is used to build the models, and the validation data is used to test how well models generalize to new data. The models and their evaluation are displayed in the *Eureqa* graphical user interface, as shown in Fig. 11.6. Obviously there is no guaranty that a model with a good correlation exists or can be found for a given data set. However, the approach and this tool provide a good way to explore the collected data and identify possible relationships that can be investigated further later on, such as

$$\text{errors} = A * \text{frustration} + B$$

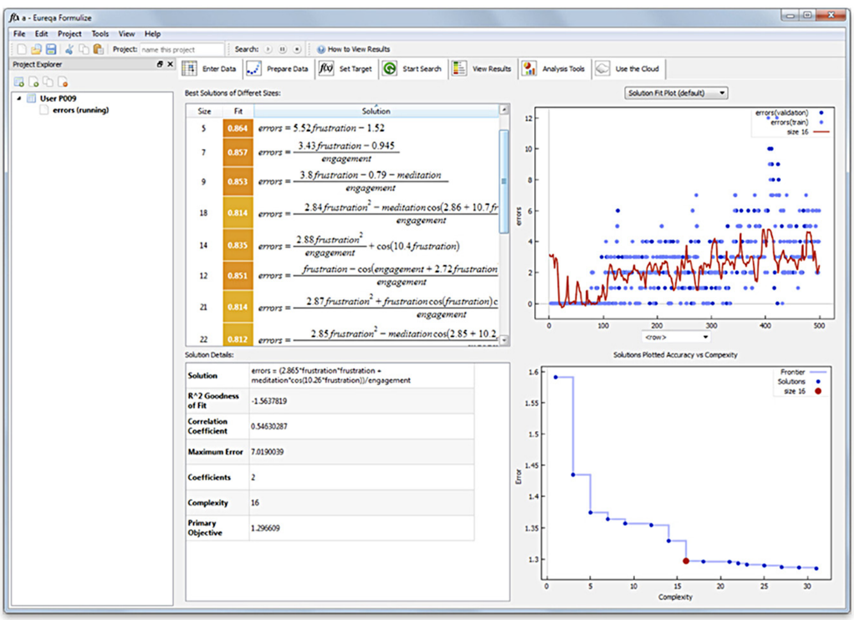

FIGURE 11.6 **Eureqa tool during the analysis of a dataset investigating the relationship between affect measurements and the error-making tendency while playing a video game.**

or

$$\text{errors} = (A * \text{frustration} - B * \text{meditation} - C)/\text{engagement}$$

The letters A, B, and C represent constant values weighting the variables frustration, meditation, and engagement. This nonconclusive exploration of mistakes made by a player of the Guitar Hero video game suggest a proportional relationship between the tendency to make a mistake and the levels of frustration and meditation, but an inversely proportional one with the level of engagement.

Classification Example

Weka[e] is a tool that implements a collection of machine learning algorithms for data mining tasks; these explore data composition and relationships and extract useful information by means of clustering and classification approaches (Hall et al., 2009). To exemplify the classification approach, let us consider a research study run to predict the level of difficulty for solving four puzzles. Each puzzle has a different difficulty level ranging from easy to hard. The Tobii Eye Tracker and the Emotiv headset were used to collect data from 44 users solving each puzzle; both raw EEG data and emotional constructs from the Emotiv headset were taken into account. With that data, the intention was to construct a prediction model to show if observed sensor data can provide information about the puzzle difficulty level. A model was created to understand and predict cognitive load (level of difficulty experienced in solving a problem), where sensor readings are the independent variables and puzzle difficulty level is the target variable (Joseph, 2013). Since both raw data and constructs were taken into account, this example allows the description of the complete set of steps presented before, as follows:

1. *Preprocessing.* For the eye-tracking data and Emotiv emotional constructs, no preprocessing was performed because it was assumed that they were already derived values and they are not as sensitive as EEG raw data. For the raw EEG data, median filtering was applied to reduce signal noise.
2. *Feature extraction.* For Emotiv constructs and pupil data, extracted features consisted of measures of variability, including variance, the minimum and maximum values, and skewness and kurtosis, as well as measures of central tendency, including mean and median. These features provide information about the characteristics of the signal distribution. For the raw EEG data, feature extraction consisted of transforming raw data into wave-frequency domains using a fast Fourier transform (FFT) algorithm and by band-pass filtering the signals within specific frequency ranges for alpha (8–12 Hz), beta

[e]http://www.cs.waikato.ac.nz/ml/weka/

TABLE 11.10 Features Calculated From Emotiv Affective Constructs and Their *P* Values for Predicting Puzzle Difficulty Level

Feature	Importance	*P* value
Long-Term Excitement_min	100%	<0.001
Long-Term Excitement_variance	98.13%	<0.001
Meditation_kurtosis	86.86%	<0.001
Engagement/Boredom_skew	60.82%	0.012
Engagement/Boredom_min	56.68%	<0.001
Frustration_variance	42.81%	0.004
Engagement/Boredom_mean	34.66%	<0.001
Engagement/Boredom_max	32.96%	<0.001
Short-Term Excitement_variance	25.66%	<0.001
Engagement/Boredom_variance	9.09%	<0.001

(13–30 Hz), and theta (4–7 Hz) waves. EEG alpha activity reflects attention demands, beta activity reflects emotional and cognitive processes, and theta activity has been related to consciousness.

3. *Feature selection.* From the 30 features extracted from the 5 Emotiv constructs and pupil data (5 measures of variability for each), 10 features were selected according to the process described by Tuv et al. (2009). Those features have *P* values less than 0.05; a small *P* value means corresponding variable values change significantly for different puzzle difficulty levels. Features, importance, and *P* values for each feature are listed in Table 11.10. Alpha, beta, and theta waves were retained as EEG features.
4. *Data mining algorithm.* The mining approach used for searching patterns that predict the puzzle being solved, that is, differentiating the puzzle type (difficult level), was classification with trees. A first approach was to train a decision tree using the *Weka* tool. The decision tree algorithm used was J48 (Quinlan, 1993). Fig. 11.7 shows a sample tree (not actual results) to exemplify a decision tree using only the features calculated from Emotiv constructs shown in Table 11.10. For instance, Fig. 11.7 shows that long-term excitement has taken a minimum value greater than 0.467 for 25 subjects while solving puzzle 4 (a higher level of difficulty), indicated by the branches on the right. Looking at the branches on the left, minimum long-term excitement values are less than 0.169 for 39 subjects while solving puzzle 1 (lower level of difficulty).

Trees are useful for understanding the patterns, but for some objectives a single tree suffers from problems with generalization—they tend to overfit. To avoid this, an approach with multiple trees can be introduced

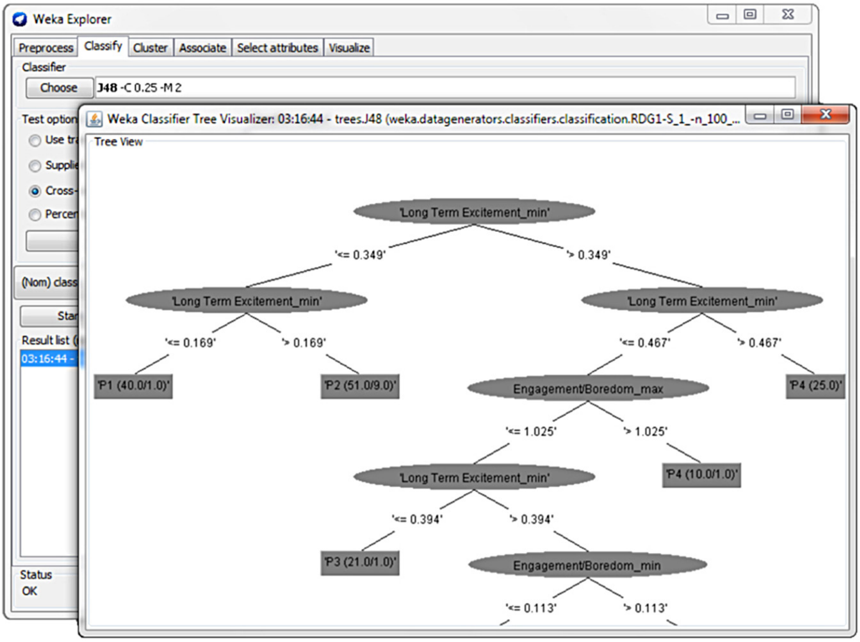

FIGURE 11.7 **Weka tool during the analysis of a dataset investigating the relationship between affect measurements and the level of difficulty solving puzzles – a J48 decision tree generated by the tool.**

in a randomized way; this approach is called Random Forest (Liaw and Wiener, 2002). Using trees for interpretation purposes and Random Forest for prediction is an option that we have explored with outstanding results.

SUMMARY

This roadmap aims to provide the reader with a better understanding of the technologies and approaches available for affect measurement, as well as an insight into how to process and analyze the data acquired. Some final thoughts are presented here along with a summary of each section.

The first section describes how a wide range of diverse approaches and technologies are being used and explored to recognize and quantify affective states. Emotions are complex constructs fused with physiological and electrical activity in our bodies. Sensing physiological and electrical activity, looking for changes, and finding patterns in them enables us to identify the affective state that is behind those changes.

The capabilities of the various approaches and technologies surveyed in the first section are summarized in Table 11.11. The data gathered by

TABLE 11.11 Summary of Approaches and Technologies for Affect Measurement

Device	Software	Input	Output (raw data and constructs)	Rate
Emotiv EPOC headset	Emotiv SDK	Brain waves	EEG activity: reported in 14 channels: AF3, F7, F3, FC5, T7, P7, O1, O2, P8, T8, FC6, F4, F8, and AF4 Face activity: blink, wink (left and right), look (left and right), raise brow, furrow brow, smile, clench, smirk (left and right), and laugh Emotions: excitement, engagement/boredom, meditation, and frustration	128 Hz (raw data) and 8 Hz (constructs)
Standard webcam	MIT Media Lab Mind-Reader	Facial expressions	Emotions: agreement, concentration, disagreement, interest, thought, and unsureness	10 Hz
Tobii Eye Tracker	Tobii SDK	Eye movements and pupil size	Gaze points (x, y), fixation, and pupil dilation	60 Hz
Skin conductance sensor	USB driver	Skin conductivity	Arousal	2 Hz
Heart rate sensor	USB driver	Heart rate	Arousal	Variable
Pressure sensor	USB driver	Pressure	One pressure value per sensor allocated into the input/control device, which can be related to frustration	6 Hz
Posture sensor	USB driver	Pressure	Pressure values in the back and the seat cushions (in the right, middle, and left zones) of a chair that can be related to interest	6 Hz
NA	Synesketch	Text	Happiness, sadness, anger, fear, disgust, and surprise	Variable

them is diverse and they can be used individually or in combination, depending on the goal of the study or the research questions being investigated. All technologies produce raw data but only some have a built-in capacity to infer constructs. In addition to each device's technical specifications, cost is also a key factor when deciding on a device or approach appropriate to a project. BCI devices, particularly the ABM B-Alert X-Series systems, are costlier in comparison to other systems discussed here. These are closely followed by eye-tracking systems embedded in glasses, the mobile and embedded eye-tracking systems, and the Emotiv EPOC headsets. For the face-based recognition systems the cost depends not on the hardware but the software, and there are a number options on the market. The least expensive systems are the skin conductance sensors, the pressure sensors, and the posture sensors, because these can feasibly be assembled in-house following blueprints available from many sources, such as those referenced above.

The second section describes the first steps in dealing with the data obtained, involving sampling, filtering, and integration. Diverse sources have diverse sampling rates; the first challenge is to synchronize these. A second challenge involves noise reduction or filtering, meaning the elimination of samples with no useful values (out of range as a result of faults in the hardware or the environment, low battery power, interference, glare, target out of the camera's line of sight, etc.). Last but not least, integration is the combination of data from diverse sources to improve the inference of affective states.

Finally, the last section explains an overall strategy for data analysis that consists of four steps: feature extraction, feature selection, data mining, and interpretation and evaluation. This section shows two examples with freely available tools, one using a regression and the other a classification approach, respectively dealing with the prediction of mistakes in videogames and puzzle solving. None of these examples are intended to present research results or conclusions. Instead, they provide an exemplary exploration of experimental date for demonstration purposes.

Acknowledgements

Our research in affect recognition and affect-driven adaptive systems is supported by the Office of Naval Research under Grants N00014-10-1-0143 and N00014-13-1-0438, awarded to Dr. Robert Atkinson. We thank Robert Christopherson, iMotion support specialist, for his advice in the initial stage of this manuscript. We would also like to thank our colleagues in the iLUX and ANGLE labs at Arizona State University for reading the manuscript and sharing their thoughts and responses, Irfan Kula in particular, for sharing his expertise and knowledge.

References

American Electroencephalographic Society, 1990. Guidelines for Standard Electrode Position Nomenclature. American Electroencephalographic Society.

Appelhans, B.M., Luecken, L.J., 2006. Heart rate variability as an index of regulated emotional responding. Rev. Gen. Psychol. 10 (3), 229–240, http://doi.org/10.1037/1089-2680.10.3.229.

Bradley, M.M., Miccoli, L., Escrig, M.A., Lang, P.J., 2008. The pupil as a measure of emotional arousal and autonomic activation. Psychophysiology 45 (4), 602–607, http://doi.org/10.1111/j.1469-8986.2008.00654.x..

Calvo, R.A., D'Mello, S.K., 2010. Affect detection: an interdisciplinary review of models, methods, and their applications. Trans. Affect. Comput. 1 (1), 18–37, http://doi.org/10.1109/T-AFFC.2010.1..

Cohn, J.F., De la Torre, F., 2014. Automated face analysis for affective computing. In: Calvo, R., D'Mello S., Gratch J., Kappas A. (Eds.), The Oxford Handbook of Affective Computing. Oxford University Press, USA.

Cooper, D.G., Arroyo, I., Woolf, B.P., Muldner, K., Burleson, W., Christopherson, R.M., 2009. Sensors model student self-concept in the classroom. Presented at the Proceedings of the 17th International Conference on User Modeling. Adaptation, and Personalization. Springer-Verlag, Berlin, Heidelberg, pp. 30–41. http://doi.org/10.1007/978-3-642-02247-0_6.

Dubcakova, R., 2011. Eureqa: software review. GPEM 12 (2), 173–178, http://doi.org/10.1007/s10710-010-9124-z..

Ekman, P., 1992. Are there basic emotions? Psychol. Rev. 99 (3), 550–553.

Ekman, P., Freisen, W.V., Ancoli, S., 1980. Facial signs of emotional experience. J. Pers. Soc. Psychol. 39 (6), 1125–1134, http://doi.org/10.1037/h0077722..

Gilroy, S.W., Cavazza, M., Benayoun, M., 2009. Using affective trajectories to describe states of flow in interactive art. Presented at the Proceedings of the International Conference on Advances in Computer Entertainment Technology. ACM, New York, NY, USA, pp. 165–172.

Hall, M.A., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H., 2009. The WEKA data mining software: an update. SIGKDD Explor. 11 (1), 10–18, http://doi.org/10.1145/1656274.1656278..

Joseph, S., 2013. Measuring cognitive load: a comparison of self-report and physiological methods. ProQuest Dissertations and Theses. Arizona State University, Ann Arbor.

Kaliouby, El, R., Robinson, P., 2005. Generalization of a vision-based computational model of mind-reading. Presented at the Proceedings of the Third International Conference on Affective Computing and Intelligent Interaction. Springer, pp. 582–589.

Krcadinac, U., Pasquier, P., Jovanovic, J., Devedzic, V., 2013. Synesketch: an open source library for sentence-based emotion recognition. Trans. Affect. Comput. 4 (3), 312–325, http://doi.org/10.1109/T-AFFC.2013.18..

Liaw, A., Wiener, M., 2002. Classification and Regression by randomForest. R News 2 (3), 18–22.

Little, R.J., Rubin, D.B., 2002. Statistical Analysis with Missing Data, second ed. Wiley, New York, NY, USA.

Manolakis, D.G., Ingle, V.K., Kogon, S.M., 2000. Statistical and Adaptive Signal Processing: Spectral Estimation, Signal Modeling, Adaptive Filtering, and Array Processing. McGraw-Hill, Boston, MA, USA.

Mehrabian, A., 1996. Pleasure-arousal-dominance: a general framework for describing and measuring individual differences in temperament. Curr. Psychol. 14 (4), 261–292.

Mota, S., Picard, R.W., 2003. Automated posture analysis for detecting learner's interest level. Presented at the Proceedings of the 2003 Workshop on Computer Vision and Pattern Recognition, vol. 5. pp. 49–55.

Novak, D., Mihelj, M., Munih, M., 2012. A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing. Interact. Comput. 24 (3), 154–172, http://doi.org/10.1016/j.intcom.2012.04.003..

Picard, R.W., 1997. Affective Computing, first ed. MIT Press, Cambridge, MA, USA.

Picard, R.W., 2010. Affective computing: from laughter to IEEE. Trans. Affect. Comput. 1 (1), 11–17, http://doi.org/10.1109/T-AFFC.2010.10..

Qi, Y., Picard, R.W., 2002. Context-sensitive bayesian classifiers and application to mouse pressure pattern classification. Presented at the Proceedings of the 16th International Conference on Pattern Recognition, vol. 3. pp. 448–451. http://doi.org/10.1109/ICPR.2002.1047973.

Quinlan, J.R., 1993. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, San Francisco, CA, USA.

Russell, J., 1980. A circumplex model of affect. J. Pers. Soc. Psychol. 39 (6), 1161–1178.

Strauss, M., Reynolds, C., Hughes, S., Park, K., McDarby, G., Picard, R.W., 2005. The hand-wave bluetooth skin conductance sensor. In: International Conference on Affective Computing and Intelligent Interaction. Springer, Berlin, Heidelberg, pp. 699-706.

Tuv, E., Borisov, A., Runger, G., Torkkola, K., 2009. Feature selection with ensembles, artificial variables, and redundancy elimination. J. Mach. Learn. Res. 10, 1341–1366.