Christian Jutten | Editor-in-Chief | Christian.Jutten@grenoble-inp.fr

Affects and Emotions in *IEEE Signal Processing Magazine*

IEEE Signal Processing Magazine (*SPM*) is a fully edited magazine. This means that, after the review process, the final files provided by the authors are processed by a team of editors under the supervision and the responsibility of the magazine's managing editor, Jessica Welsh. The contributions of Jessica and the editorial and production teams are very important, since they transform all the final files, figures, and images, into aesthetically pleasing articles that can have colorful sidebars, catchy, short titles, etc. Even if text is not changed substantially, the articles become much more visually attractive after they have been edited. These changes are unanimously appreciated by the authors, who usually cannot recognize their final articles!

Unsung heroes

I cite Prof. Min Wu, a previous *SPM* editor-in-chief, who said in her final editorial [1]:

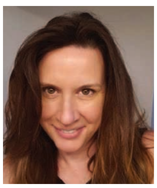
> "A number of unsung heroes, whom ordinary readers may not have seen or known, contributed to the success of our magazine. Managing Editor Jessica Welsh and the IEEE Magazines Department production team are a driving force in interacting with authors and creating a professional look and feel for the articles. In addition, Senior Art Director Janet Dudar and Associate Art Director Gail Schnitzer help bring eye-catching artistic elements to each issue of the magazine."

Since the beginning of my editor-in-chief term in January 2021, Jessica and I have had regular email exchanges (more that 1,000 in eight months!), and I appreciated the simple, constructive, and efficient interactions, that we have had together. Jessica has worked in various capacities on our magazine for more than 12 years, and the November issue is her last issue as *SPM* managing editor. In fact, *SPM* will be transitioning to a new staff editor due to some overall journal and magazine assignment changes. Jessica will continue to work with IEEE on various publications.

However, *SPM* will now be managed by Sharon Turk, who has been with IEEE for a little more than 10 years. Over the last three months, Jessica and Sharon have worked together to prepare *SPM*'s September and November issues.

I would not like that such heroes remain unsung. Thank you, Jessica for everything. It was a great pleasure to work with you. I hope the best for you. And welcome, Sharon!

Jessica Welsh



I have been with IEEE since 2006 and have worked on a variety of publications over the last 15 years. As the managing editor of *IEEE Signal Processing Magazine* for more than 12 years, I have had the opportunity to work with many wonderful editors-in-chief, as well as area editors, Society presidents, and Society volunteers and staff. I received my B.A. degree in journalism from the College of New Jersey (also known as TCNJ), in Ewing, New Jersey.

Sharon Turk

I am a journals production manager in IEEE Publishing Operations. I just celebrated my 10-year anniversary with IEEE. I work on a variety of publications that includes both transactions and magazines in the areas of wireless communications, antennas and propagation, aerospace and electronic systems and others. I received my bachelor of arts degree in English from Rutgers, the State University of New Jersey.

In this issue

SPM's November issue is a special issue (SI) on signal processing for affective computing. You will find the context and potential applications of affective computing as well as a short presentation of the nine articles of the SI in the guest editorial on page 9. Affective computing especially is interested to recognize, synthesize, and react to emotions using different modalities: speech and text recordings, images and videos for extracting face expressions and gestures, physiological data, etc. For most readers, affective computing probably

Digital Object Identifier 10.1109/MSP.2021.3096153
Date of current version: 27 October 2021

seems far from their preferred topics, but I strongly encourage them to read the articles, since I found them very inspiring for many domains in signal processing or, more generally, in data science. Even if the articles are focused on various aspects of affective computing, they are based on different recording modalities and target different objectives, and they all consider very carefully the data, and discuss in depth many issues related to data.

Even if the articles are focused on various aspects of affective computing, they are based on different recording modalities and target different objectives, and they all consider very carefully the data, and discuss in depth many issues related to data.

Based on face recordings, Ekman [2] proposed six basic emotion categories of facial expressions: happiness, surprise, fear, disgust, anger, and sadness. On the face as well as in speech recordings, from our own experience, we easily understand that the features associated to these categories of emotion are strongly variable and interdependent, e.g. of the gender, personality, social role, health condition, language, countries, and culture, etc. The features are also strongly dependent of the image quality, of the noise and environment variability: features extracted on real-world recordings used in the inference step are usually very different than features extracted on lab-controlled recordings.

The data sets are usually annotated, which is mandatory for supervised learning methods. But, due to the perception variability between annotators, there is no ground truth: annotations are also strongly dependent of the gender, the culture, personality, etc.

In fact, following Booth et al. (see Figure 3 in their article on page 88 in this issue) one can view the complete process in an information theory framework: the transmitter sends signals (facial expression, speech, gestures) coding "emotions," signals are then corrupted through a noisy channel, the receiver decodes the noisy signals. But, in affective computing, the transmitter and the receiver don't often share the same code!

So, in all the articles of the special issues, problems of data representativeness, like bias and fairness, are ubiquitous. Concepts for measuring bias and fairness, for detecting attributes sensitive to bias and fairness, are discussed in depth. For avoiding bias and variability in annotations, the authors proposed to avoid supervised learning, and suggest unsupervised, semisupervised, or self-supervised learning. For considering signal variability, present in facial/speech expressions (transmitter level), in the perception (receiver level) or due to noise, which can change the feature distribution, solutions like transfer learning or domain adaptation (see [3] for a review) are suggested.

In all of these articles, statistical learning methods are widely used, and especially, the authors recognized the interest of deep learning, but not as a black box: conversely, they point out the importance of explainability, robustness, and replicability.

Finally, scientists working on affective computing are faced with ethical issues. Of course, their ideas, methods, and results are designed to be used for useful applications including health, better human–machine interactions, etc., but could be so easily diverted to malicious uses!

So, I believe that discussions, solutions, and methods presented in all of these articles are very inspiring for many current problems in signal processing. Finally, I have two timely questions and, since I was sure that readers could have the same concerns, I consulted with Prof. Björn Schuller, the lead guest editor of this SI, and I warmly thank him for his answers:

Q: Considering the trend of fakes, are some methods being designed for detecting fake emotions?

A: Yes, "regulated" emotions (e.g. "posed smile", etc.) are present in data which have been used in Interspeech ComParE challenges that I organized. And I also saw some papers speaking about "fake" emotions, e.g. in an international Challenge at ICCV 2017 [4]. However, it is hardly addressed in comparison to recognition of the perceived emotion as by human observers (the standard case) and there is very little data on faked ones.

Q: In the current COVID-19 context, I would be curious to know: What is the performance of the different methods of emotion recognition when a mask covers a large part of the face, and what efficient solutions the authors could propose for considering such change in the face recordings?

A: We showed in the same challenge series at Interspeech 2020 ComParE that one can recognize automatically from audio if someone is wearing a facial mask with the same methods we would usually recognize emotions, but we have not analyzed the impact on emotion recognition explicitly. For the face, this is ultimately facial expression recognition under occlusion, which has repeatedly been addressed. I found, e.g., a paper addressing specifically COVID-19 facial masks [5].

References

[1] M. Wu, "Signals and signal processing: The invisibles and the everlastings [From the Editor]," *IEEE Signal Process. Mag.*, vol. 34, no. 6, pp. 4–7, Nov. 2017. doi: 10.1109/MSP.2017.2750306.

[2] R. Ekman, *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS).* New York: Oxford Univ. Press, 1997.

[3] V. M. Patel, R. Gopalan, R. Li, and R. Chellappa, "Visual domain adaptation: A survey of recent advances," *IEEE Signal Process. Mag.*, vol. 32, no. 3, pp. 53–69, May 2015. doi: 10.1109/MSP.2014.234705.

[4] J. Wan et al., "Results and analysis of ChaLearn LAP multi-modal isolated and continuous gesture recognition, and real versus fake expressed emotions challenges," in *Proc. IEEE Int. Conf. Comput. Vision Workshops (ICCVW)*, 2017, pp. 3189–3197. doi: 10.1109/ICCVW.2017.377.

[5] B. Yang, W. Jianming and G. Hattori, "Face mask aware robust facial expression recognition during the COVID-19 pandemic," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2021, pp. 240–244. doi: 10.1109/ICIP42928.2021.9506047.

SP