



EmotionML Position Paper

Emotions in Consumer and Marketing Research

Tim Llewellynn (tim.llewellynn@nviso.ch), Matteo Sorci (matteo.sorci@epfl.ch)

nViso Sàrl is a software company that empowers brands, agencies, and consumer businesses to understand human behavior through automated facial image analysis. A young start-up from the EPFL signal-processing laboratory based in Switzerland, it has developed nonintrusive technologies for deciphering emotions by automatically detecting human facial expressions, head gestures, and eye movements. It is a timely innovation for a range of consumer-facing industries as emotions are the driving force of life, fundamental to human experience, influencing perception, and everyday tasks such as communication and decision-making. However, they have largely been ignored by companies, partly because emotions are hard to measure, difficult to quantify, and their affect has been often misunderstood. nViso has taken steps to overcome these problems through a patented and proprietary image and video analysis platform with easily understandable and quantifiable outputs.

Emotions and feelings have long been recognized as important factors in consumption and consumer decision making. The understanding of emotions and the measurement of emotions in consumer research has throughout time depended largely on contributions from more grounded disciplines e.g. psychology and sociology, and the increased focus on emotions in other disciplines has given ground for increasing attention in consumer research. The recognition of the importance of emotion in decision-making, not least due to findings in cognitive neuroscience, has escalated the attention. EmotionML gives an unique opportunity to partly unify what is currently a long list of approaches to measuring emotions in consumer and marketing research. However it also raises many questions on complex and difficult issues, many of which are still areas of active research. We wish to present a specific use case of using emotions to assess the impact of online videos and how EmotionML can be practically used in this context and to open several discussion on open questions :

1. How should emotion vocabularies be defined? Should this be driven by academic theory, business needs, or what we can reliably measure? How should descriptions be managed?
2. Context gives meaning. Should the nature and context of how emotions are experienced, expressed, or captured be given? If so, how?
3. How can reliability and confidence of captured data be expressed? Who gives meaning to confidence=0.5 and how is confidence defined?

nViso Emotion Profiling based on Facial Expressions

With the mass transition to online social media, companies and consumers have lost an important part of communication: our instant expression of feelings and emotions. Today, we are doing more and more faceless, and therefore expressionless communicating. E-mail, internet, text-messaging, mobile telephones all expedite language based communication at lightning speed. And businesses are relying on it increasingly as a medium for complex communications, negotiation, and decision making. But in evolutionary terms, language is a relatively new addition and has its limitations.

Our faces can express things that are difficult to put into words. Expressions can communicate emotions faster, more subtly and more effectively than words ever can, which is why facial expressions remain crucial for humans as social animals. In order to tackle this problem and move towards natural human computer interaction, computers should capture, mimic and reproduce human perceptions, where facial expressions clearly play a central role.

In working towards this goal, nViso has developed a scientific means to accurately translate and quantify the human perception of facial expressions through what is termed “emotion profiling” that is free from cognitive and cross-cultural bias that plague existing approaches. It uses state-of-the-art video analysis and data mining techniques to accurately map and track small facial muscle movements and decode and interpret these movements based on facial coding, that was originated by Charles Darwin, refined by Dr. Paul Ekman, and being brought into daily business practice by nViso.

The system developed by nViso gives a psycho-physiological measurement compared to the traditional cognitive testing methods in measuring emotional responses of consumers. It enjoys three major advantages:

- **Universality** - expressions aren't socialized, they're "hardwired" into our brains; as a result, even a person born blind has the same facial expressions and children as young as 1.5 years-of-age already exhibit all the core emotions.
- **Spontaneity** - the face is the only place in the body where the muscles attach right to the skin, resulting in real-time data. The muscles can be controlled directly from the sub-conscious brain.
- **Abundance** - human beings have more facial muscles than any other species on the planet, ensuring a wealth of information.

One of the challenges in using the face to measure non-verbal responses lies in building a framework and system to accurately detect correction facial expressions. Several approaches to this task have been proposed over the last decade by several researchers (Tian[3], Cohen[4], Pantic[5], Hu[6], Bartlett[7], Zhang[8]). They generally suffer from several important shortcomings, namely :

- They concentrate on “what is perceived” while partially neglecting the other important issue for face-to-face interaction, that is “who is perceiving”.
- They do not take into account the subjectivity in the perception of an expression, because they rely on the judgment of only a few experts.
- They are often unable to interpret and reuse the knowledge acquired by the automatic system, because of the “black-box” nature of the used framework.

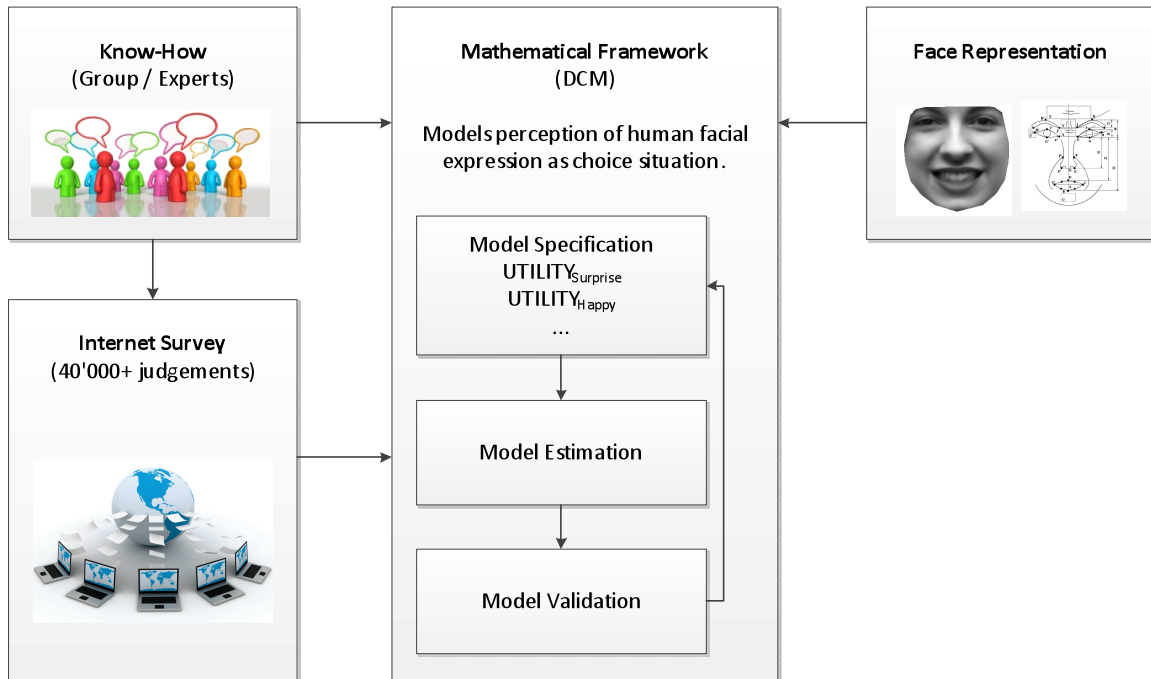


Figure 1 : nViso's unique framework in modeling human perception of facial expressions.

In order to overcome these shortcomings, a new approach used by nViso is inspired by the pioneering work of Dr. Matteo Sorci [1] [2]. The novelty of this approach is that it defines new modeling algorithms capable of taking into account the heterogeneity of human perception of facial expressions and flexible to deal with cross cultural issues. The perception process is modeled as a choice process where individuals have to choose, based on their own perception, among the set of primary emotions. As such, an econometric approach has been used based on Discrete Choice Models (DCM), a family of econometric models designed to forecast the behavior of individuals in choice situations. Relying on judgment from a heterogeneous set of many thousands of individuals from all around the world with a variety of cultural backgrounds, ages and gender, belonging to different ethnic groups and collected over the last 6 years. The estimated model has vastly outperformed the state-of-the-art approaches and revealed the potential extension to learn new non-verbal behaviors.

References

- [1] Capturing Human Perception of Facial Expressions by Discrete Choice Modelling, Sorci, Matteo ; Antonini, Gianluca ; Cruz Mota, Javier ; Robin, Thomas ; Bierlaire, Michel ; Thiran, Jean-Philippe, In: Choice Modelling: The State-of-the-Art and the State-of-Practice, 2010, p. 101-136 Emerald Group Publishing Limited, 2010.
- [2] Modelling human perception of static facial expressions, Sorci, Matteo ; Antonini, Gianluca ; Cruz Mota, Javier ; Robin, Thomas ; Bierlaire, Michel ; Thiran, Jean-Philippe, In: Image and Vision Computing, vol. 28, num. 5, 2010, p. 790-806, Elsevier, 2010.
- [3] Evaluation of gabor-wavelet-based facial action unit recognition in image sequences of increasing complexity, Y. li Tian, T. Kanade, J.F. Cohn, in: Proceedings of the 5th IEEE International Conference on Automatic Face and Gesture Recognition, 2002, pp. 229-234.

[4] Facial expression recognition from video sequences: temporal and static modeling, I. Cohen, N. Sebe, L. Chen, A. Garg, T.S. Huang, *Computer Vision and Image Understanding* (10) (2003) 160–187.

[5] An expert system for recognition of facial actions and their intensity, M. Pantic, L.J.M. Rothkrantz, in: *National Conference on Artificial Intelligence (AAAI)*, 2000, pp. 1026–1033.

[6] Manifold based analysis of facial expression, C. Hu, Y. Chang, R. Feris, M. Turk, *CVPRW '04: Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'04)*, vol. 5, IEEE Computer Society, Washington, DC, USA, 2004, p. 81.

[7] Recognizing facial expression: machine learning and application to spontaneous behavior, M.S. Bartlett, G. Littlewort, M. Frank, C. Lainscsek, I. Fasel, J. Movellan, in: *Computer Vision and Pattern Recognition, 2005, CVPR 2005, IEEE Computer Society Conference on*, vol. 2, 2005, pp. 568–573.

[8] Active and dynamic information fusion for facial expression understanding from image sequences, Y. Zhang, Q. Ji, *Transactions on Pattern Analysis and Machine Intelligence* 27 (5) (2005) 699–714.