Review of BSP lecture Digitalising the Sense of Smell by Alex Wiltschko and BSF lecture Modern Strategies and Techniques of Analysis of the Volatile Fraction as a Tool for Food Control and Characterisation by Carlo Bianchi

Dr Ali Green

Analysis, AI and the Aroma Trades – what does the future hold?

I had the pleasure of attending two fascinating online lectures in May, one from the British Society of Perfumers that investigated if AI could be used to develop some kind of artificial smelling tool or 'osmoscope' and the second exploring how food chemistry, analytical techniques & large data sets in the field of metabolomics (the study of metabolites) could be used to assess quality of specific flavour ingredients. Both speakers discussed science (and some maths) that was mind-blowingly complex (at least for me!) so I will not be attempting to replicate the details here, rather I will attempt to give an overview and then provide links to academic publications that will provide the details for those who would like to explore this ground-breaking field in more depth.

Alex Wiltschko of Google Brain shared some of the research his team have been working on exploring machine learning for olfaction. Initially he wanted to ensure that novices (like me) were able to grasp the meaning of key terms like artificial intelligence (AI) and machine learning (ML): there is no agreed definition of AI but most AI applications are based on ML defined as 'techniques that let software learn behaviour from example data rather than rules defined by the programmer'. The four main techniques of ML are: -

- Classification
- Prediction
- Generation
- Language understanding

These are usually employed as part of a problem-solving strategy and are in use across numerous fields from agriculture to healthcare. Success in ML depends upon several key factors: good data, good tools and sound research, trained people and centralised resources (since it can be prohibitively expensive to do all the jobs in house). The latest research on effective ML shows that smaller data sets prove more accurate than large ones (although this is not always possible in some fields of research that have a wider scope) and very specific and clear use of language that avoids ambiguities and muddy thinking. It's important to remember that ML is always a collaboration between algorithms[1] and humans. When ML is employed effectively, it has proved invaluable in machine perception (robots/machines with video and audio recognition e.g. computerised pathology analysis in medicine), generative models (computerised predictions used widely in the financial sector) and unsupervised and reinforcement learning.

It would be tempting to assume that we could approach digitalising the sense of smell in the same way that computerised vision and audio capabilities have been developed. However, more than 200 years of research lies behind recent developments with the senses of sight and hearing being pretty much understood by the middle of the last century with early vision and audio theories going right back to Aristotle and the Greek Atomist philosophers! Understanding of the senses of smell and taste has been far more elusive and challenging, with a number of theories still being explored. It is fair to say that one of the main reasons that this is a challenge is that, unlike the science of sight, where photons and their interaction with the eye is fully grasped and can therefore be replicated in a machine, the interaction between scent and taste molecules and our olfactory system is only beginning to be understood. There does not seem to be an obvious relation between molecular structure and aroma that could be plugged into a digital device along the lines of "if it has a specific functional group, it will always smell of lilies" since similar molecules don't always smell the same and disparate ones sometimes do! The mysteries of aroma chemistry therefore make this more of a challenge for those wanting to develop an 'osmoscope'.

The team at google brain were determined to discover some kind of pattern in the data and so painstakingly recorded the structural features and aroma qualities of 5000 molecules. They recorded these on a graph neural network (GNN)[2]** to see if there were any meaningful correlations between structure and aroma[3]. Inside the NN, around 60 dimensional embeddings were organised intuitively into separate categories that fitted into broad odour descriptors – this at some times seems to reflect the molecule structure but not always, as some families of molecules were distributed across numerous activity centres. An interesting question is thus why are they forming these GNN patterns since nature doesn't generally evolve in a purely random way, there must be some kind of perhaps evolutionary need or reason. What is clear is that this reason cannot be nutritional because nearly all molecules of nutritional value are odourless apart from short-chain fatty acids which actually smell rancid! Although there are undoubted therapeutic qualities for some of nature's fragrant plants such as lavender, many other 'useful' plants have no scent at all like *plantago lanceolata* (ribwort plantain). So what is going on?

The Google Brain team have posited that olfaction has evolved to detect the chemical processes that create or result from the molecules that affect our survival and are correlated to positions in the carbon cycle. The plots on the GNN appear to correlate to fruiting photosynthesis (grape family), flowering photosynthesis (lily family), animal respiration (musk family), combustion (toasted family), putrefaction and decay (mushroom family) and fermentation (rum family). The hypothesis then examined various aromas and what they might signal to a sentient being:

- Rotting meat in itself doesn't make us sick – it's the enterotoxins the putrefying bacteria secrete in the gut. Thus, decarboxylation of amino acids into compounds charmingly named putrescine and cadaverine only create a revolting but harmless smell.
- The smell of baking bread or cooking meat from Mallard reaction products doesn't in itself feed us, rather it signals the carbohydrates and amino acid precursors.
- Smells of specific places are absolutely vital to survival – ocean is not fruit, forest fire is not rot and meat isn't dirt!

So, does this interesting theory and the research of the Google Brain team bring us any closer to an 'osmoscope'? The answer is tantalisingly perhaps, but it is unlikely to be in the near future. What is clear is that if smell could be digitalised, it would bring numerous benefits to society across numerous fields – from healthcare to the environment as well as industry applications associated with QC, formulation and fragrance creation.

Alex was keen to hear from anyone in the industry who would like to contribute or comment on the team's research. He can be reached on alexbw@google.com

He also highlighted a new publication on Interpretability in Chemistry by the team: -

Sanches-Lengeling, Benjamin et al., 2020, Evaluating Attribution for Graph Neural Networks, *Advances in Neural Information Processing 33 2020*

The British Society of Flavourists brought in expert Food Chemist Carlo Bicchi of Turin University's Laboratory of Pharmaceutical Biology and Food Chemistry to discuss his team's ground-breaking work on the sense of taste and how this could be digitalised. Carlo first discussed the situation in the recent past, highlighting how there had been a traditional gap in attitude towards the sensory skills traditionally employed by perfumers (viewed by scientists as a more fluid/creative approach and thus imprecise) and the use of analytical chemistry (viewed by perfumers as rigid and unable to equal the olfactory capabilities of a living being).

In order to explain how the new union of chemistry and biology known as sensomics is helping to pioneer a new approach in flavour quality control, Carlo began with a few definitions. When examining a volatile fraction[4] analytically the first thing to note is that each fraction has a different and mutually non-comparable composition to others. each comprising its own headspace, essential oils, aromas[5], flavours[6], fragrances[7] and extracts.

A new approach, part of the field known as metabolomics, examines metabolite[8] fingerprints from specific cellular processes: according to Carlo *"the goal is a comprehensive and quantitative analysis of the largest possible array of low-molecular weight metabolites"* in biological samples. Within metabolomics, there is a subdivision that focuses solely on sensory studies known as sensomics which heralds a new and far more effective kind of analysis combining fingerprinting and profiling. Between sensomics and chemical analysis lies the new field of sensometrics, which *"tries to establish relationships between the sensory scores of a food aroma and its chemical fingerprint or profile with chemometric tools".* The aim is a Total Analysis System (T.A.S.) linking sample preparation, analysis (separation and detection) and data elaboration (chemometrics) into one process. This is encapsulated in sensomics, a process described by Carlo as "a gold standard for food sensory characterisation". It has five components: -

1. Isolation: using the SAFE process (solvent-assisted flavour extraction)
2. Screening: using AEDA tech (aroma extract dilution analysis)
3. Identification: ITs, MS spectra, odour
4. Quantitation: using SIDA (stable isotope dilution analysis)
5. OAV measure/flavour reconstruction (detecting the odour activation value - the amount at which it is detected by the senses)

One food that Carlo's team has been working on using this technology is coffee with a view towards innovation in QC. They have isolated 32 chemicals responsible for the various odour profiles associated with ground roasted coffee and have assigned useful descriptors to each flavour molecule (such as 2,3-Pentandione: buttery or 3-Isopropyl-2-methoxypyrazine: earthy/roasty) and assessed the 'strength' of the flavour by assigning each a flavour dilution factor (FD factor)[9] and also its OAV to check if there's sufficient quantity to be perceived and thus contribute to the flavour profile. Different combinations of these are found in typical beans from specific locations thus creating the unique flavour profiles for different localities and the treatments they have received (whether they have been washed or not or given other treatments)[10]. A key element of their approach has been that it is solvent free, easy to automate, reliable and representative so that it is easy to replicate. The sensomic approach integrates both analytical chemistry and sensory evaluation to reach thorough conclusions that accurately describe the subtle differences between regional coffees both from their molecular composition and sensory qualities. Thus 'fresh' notes such as acid, flowery and fruity along with 'brown' notes such as bitter, nutty, woody and spicy can be mapped according to bean type and origin very securely because of their chemical composition as well as their sensory qualities. Far more details about this complex analysis and the specifics of the chemistry can be seen in the papers detailed in the footnotes.

Carlo went on to discuss another important raw material that has been analysed in this way, cocoa, which is under pressure as a crop for a number of reasons: worldwide increase in demand and consumption and the impact of climate change on production, particularly the influence of rainfall and to some extent temperature and light conditions. A big problem in the processing of coca is the smoky 'off' flavour that can taint the end product (cocoa powder and chocolate) as a consequence of wood fires during drying and storage.Using the same approach as that used for coffee (a combination of HS-SPME-GCxGC-MS TOF analysis and sensory data) the team were able to detect a chemical 'signature' for the smoky and non-smoky samples that differed thus making it relatively straightforward to detect tainted crops at an early stage with a routine test. This means that what begins as incredibly complex chemical analysis can be translated to a routine quality control test using a far quicker mono-dimensional system with capacity for a high throughput that confirms the differences through phenolic and benzene derivatives.

Clearly the techniques discussed by Carlo are pioneering and can easily be transferred to other fragrance and flavour ingredients thereby aiding QC and ensuring trust and reducing capacity for adulteration or pollution within industry.

I would strongly recommend subscribing to the British Society of Flavourists' free publication the New Flavourist[11] and follow their Twitter feed for news on all flavour-related topics. The British Society of Perfumers[12] also have an active Twitter and Facebook feed; it's well worth checking out their events pages to join in online and in-person sessions that provide access to industry experts talking on all manner of industry subjects.

1 Algorithm definition: *"A mechanistic formula that will automatically produce an answer for each new case that comes along with no, or minimal, additional human intervention"* David Spiegelhalter The Art of Statistics (Reviewed by Tony Curtis for ICATS News in 2020)

2 Graph neural network definition: a neural network is a computing system that is based on the workings of an animal's brain using a system of nodes which replicate the neurones of a living creature. Neural networks with many layers have become known as deep-learning models. A graph neural network is a method of displaying complex NN data using a graph representation structure.

3 A detailed summary of this analysis and primary findings can be seen in this paper: https://www.semanticscholar.org/paper/Machine-Learning-for-Scent%3A-Learning-Generalizable-S%C3%A1nchez-Lengeling-Wei/19af82973b785c66c3033377eebca4513e106879

While this is Alex's blog summarising the paper in less detail: https://ai.googleblog.com/2019/10/learning-to-smell-using-deep-learning.html

4 Each volatile fraction is isolated from the overall matrix of a food material - it is a mixture of compounds released together at a specific temperature sometimes using specific solvents.

5 Aromas are usually volatile compounds consumed orthonasally and/or retronasally by the olfactory tissues in the nose

6 Flavour is the overall sensation provided by the interaction of odour and textural feeling when food is consumed.

7 Fragrances have a sweet/pleasant scent that derives from a non-food material and are orthonasally perceived,

8 Metabolites are the intermediate or end products of metabolism (chemical processes that happen within a living organism to maintain life)

9 The number of parts of solvent required to dilute the aroma extract until the aroma value is reduced to one.

10 Chemometric Modelling of Coffee Sensory Notes through Their Chemical Signatures: Potential and Limits in Defining an Analytical Tool for Quality Control, Davide Bressanello, Erica Liberto*, Chiara Cordero, Barbara Sgorbini, Patrizia Rubiolo, Gloria Pellegrino, Manuela R. Ruosi, and Carlo Bicchi,J., Agric. Food Chem. 2018, 66, 27, 7096–7109, Publication Date:June 12, 2018 https://doi.org/10.1021/acs.jafc.8b01340

E.Liberto, D. Bressanello, G. Strocci, C. Cordero, M.R. Ruosi, G. Pellegrino, C. Bicchi & B. Sgorbini, HS-SPME-MS-Enose coupled with chemometrics as an analytical decision maker to predict in-cup coffee sensory quality in routine controls: possibilities and limits, Molecules 2019, 24, 4515 https://doi.org/10.3390/molecules24244515

11 You can subscribe here: https://www.bsf.org.uk/

12 https://bsp.org.uk/