

POSTER PRESENTATION

Open Access

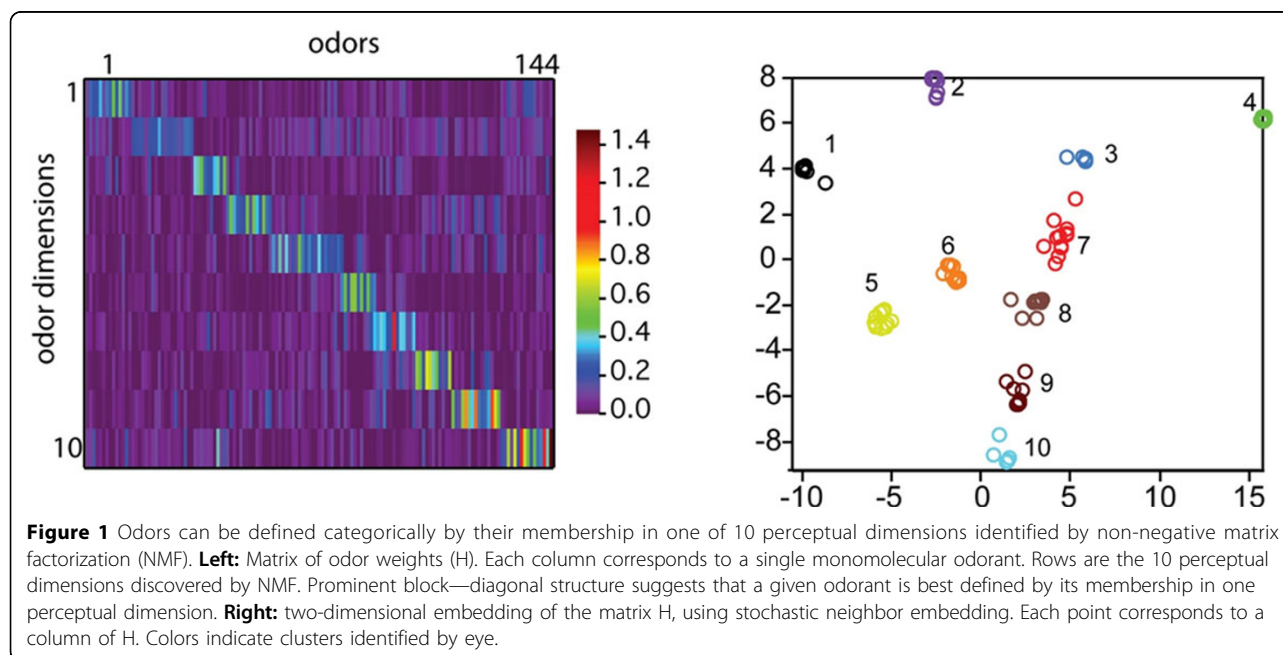
Categorical, low-dimensional decomposition of human odor space with non-negative matrix factorization

JB Castro¹, SP Quinn², A Ramanathan³, CS Chennubhotla^{4*}

From 1st International Workshop on Odor Spaces
Hannover, Germany. 4-7 September 2013

Recent studies using Principal Components Analysis (PCA) support low-dimensional models of odor space, in which one or two dimensions - with hedonic valence featuring prominently - explain most odor variability. Here we use non-negative matrix factorization (NMF) - a non-linear optimization method - to discover an alternative, reduced-dimensional representation of the Dravnieks odor database (144 odors x 146 descriptors). NMF is theoretically well-suited for these types of analyses, as odor

profiling data is inherently non-negative (e.g. descriptors either apply, or do not). We divided the dataset into training and testing halves, and found that RMSD testing error attained a minimum for subspace choice of 25, motivating this as an upper bound for odor perceptual space dimensionality. More parsimonious representations were found by comparing reconstruction errors (fraction of unexplained variance) of NMF with reconstruction errors of PCA on scrambled data (PCAsd). For subspace sizes > 10,



⁴Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA, USA

Full list of author information is available at the end of the article

NMF error was indistinguishable from PCA_sd error, indicating no gain in retaining more than 10 perceptual dimensions. As is typical of NMF basis sets, the 10 odor dimensions we obtain are sparse (only a small subset of the 146 descriptors apply), and categorical (represent a positive valued quality). Moreover, these 10 dimensions were near-orthogonal, with a mean angle of 73 degrees between all pairs of basis vectors. Investigating the distribution of odors in this 10-dimensional space, we find marked clustering (Figure 1), with each odor being well-defined by its membership in a single dimension, and to the exclusion of others. In ongoing work, we are using graph-kernel methods to define a rudimentary mapping between physicochemical features of odorants and the 10 descriptor dimensions.

Authors' details

¹Department of Psychology, Bates College, Lewiston, ME, USA. ²Joint CMU-Pitt Ph.D. Program in Computational Biology, University of Pittsburgh, Pittsburgh, PA, USA. ³Computational Data Analytics Group, Oak Ridge National Laboratory, Oak Ridge, TN, USA. ⁴Department of Computational and Systems Biology, University of Pittsburgh, Pittsburgh, PA, USA.

Published: 16 April 2014

doi:10.1186/2044-7248-3-S1-P3

Cite this article as: Castro *et al.*: Categorical, low-dimensional decomposition of human odor space with non-negative matrix factorization. *Flavour* 2014 **3**(Suppl 1):P3.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

