

Separating Style and Content

In many vision problems, we want to infer two (or more) hidden factors which interact to produce our observations. We may want to disentangle illuminant and object colors in color constancy; rendering conditions from surface shape in shape-from-shading; face identity and head pose in face recognition; or font and letter class in character recognition. We refer to these two factors generically as "style" and "content".

We introduce a general framework for analyzing the style of a multimedia signal. We assume that we can observe a training signal under several different styles. This information is often available or can be generated. We then fit those data with a bilinear model which explicitly represents the two-factor nature of the observations. The result is a modular representation of the signal which allows for independent manipulation of the two factors, style and content.

We focus on three kinds of tasks: extrapolating the style of data to unseen content classes, classifying data with known content under a novel style, and translating two sets of data, generated in different styles and with distinct content, into each other's styles. We show examples from color constancy, face pose estimation, shape-from-shading, typography and speech.

Style synthesis example

Training fonts					synthetic	actual
A	<i>A</i>	A	<i>A</i>	A	<i>A</i>	A
B	<i>B</i>	B	<i>B</i>	B	<i>B</i>	B
C	<i>C</i>	C	<i>C</i>	C	<i>C</i>	C
D	<i>D</i>	D	<i>D</i>	D	<i>D</i>	D
E	<i>E</i>	E	<i>E</i>	E	<i>E</i>	E
F	<i>F</i>	F	<i>F</i>	F	<i>F</i>	F
G	<i>G</i>	G	<i>G</i>	G	<i>G</i>	G
H	<i>H</i>	H	<i>H</i>	H	<i>H</i>	H
I	<i>I</i>	I	<i>I</i>	I	<i>I</i>	I

Style extrapolation in typography. The training data were all letters of the 5 fonts at left. The test data were all the Monaco letters except those shown at right. The synthesized Monaco letters compare well with the missing ones.

Participants

- [Bill Freeman](#) (MERL)
- Josh Tenenbaum (MIT Dept. Brain and Cognitive Sciences)

Technical Reports

J. B. Tenenbaum and W. T. Freeman, Separating style and content with bilinear models, Neural Computation

12(6), pp. 1247-1283, 2000: [MERL-TR99-04](#)

W. T. Freeman and J. B. Tenenbaum,
IEEE Conference on Computer Vision and Pattern Recognition, (CVPR '97),
Puerto Rico, U. S. A., June, 1997. (Received Outstanding Paper award).
Available as [MERL-TR96-37](#)

Separating style and content

J. B. Tenenbaum and W. T. Freeman,
in Advances in Neural Information Processing Systems 9,
M. C. Mozer, M. I. Jordan and T. Petsche, Eds., Morgan Kaufmann, San Mateo, CA, 1997.
Available as [MERL-TR96-36](#)