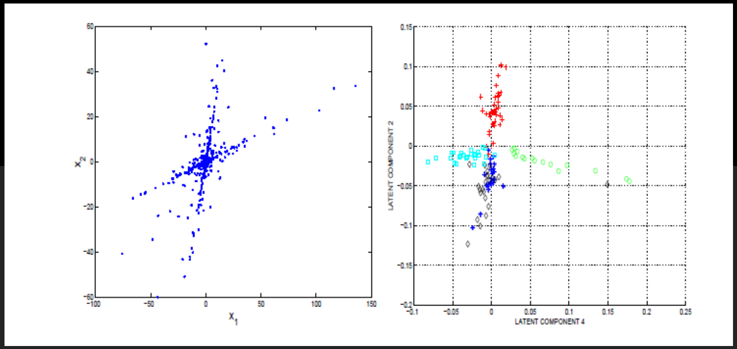
The Cognitive Components of Digital Media

...and why they matter for multi-media retrieval!

Outline



Cognitive component analysis:

- Our definition
- Motivation and related ideas in the literature
- Audio signals: Phonemes as cognitive components
- Higher order cognition: Text, indexing multi-media, social cognition
- Search engine demos

Conclusion and outlook

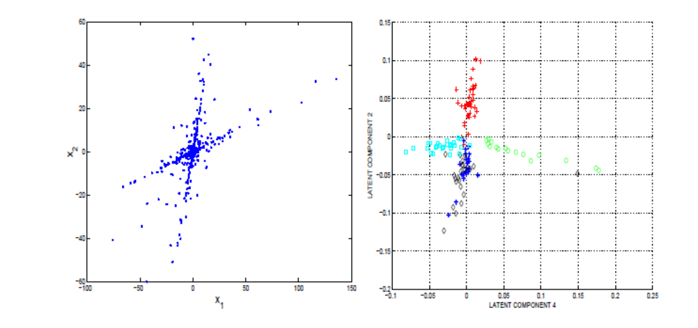
Cognitive Component Analysis

Cognitive component analysis (COCA)

- Hypothesis: Cognitive information processing is driven by statistical properties of the environment.
- The process of unsupervised grouping of data so that the resulting group structure is well-aligned with grouping based on human cognitive activity (Hansen et al., 2005).

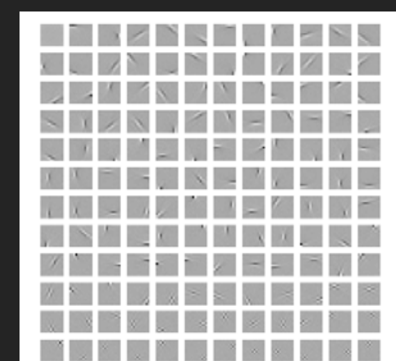
"Rational models of cognition explain human behavior as approximating optimal solutions to the computational problems posed by the environment"

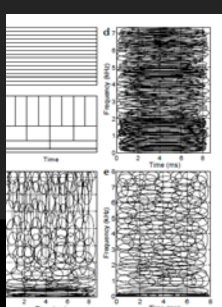
(Griffiths et al., 2007)



Cognitive compatibility as "*μ-Turing*" test....

Ecological modeling approach





Important for engineering proxies for human information processing…
Cf. efficient coding of ”context-to-action” mapping

the key question in cognition and information retrieval:
What is an object ≡ should be indexed?

Many generalizations are possible – which ones will make sense to a human?

Cognitive component analysis and the notion of *object*

The **object** is a basic notion in cognitive psychology

- E.g., Bundesens "Theory of attention" estimates number of objects in short time memory
- A pragmatic definition of an object could be: **An object is a signal source with independent behavior in a given environment**
- Cognitive component analysis is a step towards an general purpose definition of an object
- Information theory: Optimality, do brains exploit the coding advantage?

Modeling issues: We are interested in the relation between supervised and un-supervised learning. Related to the discussion of the utility of unlabeled examples in supervised learning and swift learning.

Engineering issues: Can we predict the digital media components that a human will pay attention to? -a key challenge for cognitive systems & information retrieval

Cognitive Information Processing

Cognitive Component Analysis

L.K. Hansen, P. Ahrendt, and J. Larsen. Towards cognitive component analysis. AKRR'05 - Adaptive Knowledge Representation and Reasoning. 2005,

L. Feng , L.K. Hansen. On Low-level Cognitive Components of Speech. International Conference on Computational Intelligence for Modelling, vol.2, pp 852-857, 2005.

L. Feng , L.K. Hansen. Phonemes as Short Time Cognitive Components. The 31st International Conference on Acoustics, Speech, and Signal Processing, vol.5, pp869-872, 2006.

L.K. Hansen, L. Feng. Cogito Componentiter Ergo Sum. The 6th International Conference on Independent Component Analysis and Blind Source Separation, pp 446-453, 2006.

L. Feng, L.K. Hansen. Cognitive Components of Speech at Different Time Scales. The 29th annual meeting of the Cognitive Science Society, pp 983-988, 2007.

L. Feng , L.K. Hansen. On Phonemes as Cognitive Components of Speech. The 1st IAPR Workshop on Cognitive Information Processing, pp 205-210, 2008.

L. Feng, L.K. Hansen. Is Cognitive Activity of Speech Based on Statistical Independence? The 30th annual meeting of the Cognitive Science Society, pp 1197-1202, 2008.

Emotion in song lyrics

M.K. Petersen, M. Morup, L.K. Hansen: Sparse but emotional decomposition of lyrics In Proc. LSAS 2009, International workshop on learning semantics of audio signals, Graz, Austria 2009.

M.K. Petersen, L.K. Hansen, A. Butkus: Semantic contours in tracks based on emotional tags In Proc. Computer Music Modeling and Retrieval: Genesis of Meaning in Sound and Music. Lecture Notes in Computer Science 5493:45-66 (2009).

M.K. Petersen, L.K. Hansen: Latent semantics as cognitive components In Proc. 2nd International Workshop on Cognitive Information Processing. Elba Island, Italy (2010).

M.K. Petersen, L.K. Hansen. Cognitive Semantic Networks: Emotional Verbs Throw a Tantrum but Don't Bite. Workshop on Cognitive Information Processing CIP, Baiona, Spain (2012).

Top-down attention

L.K. Hansen, S.G. Karadogan, L. Marchegiani. What to measure next to improve decision making? On top-down task driven feature saliency. 2011 IEEE Symposium on Computational Intelligence, pp. 81-87 (2011).

L. Marchegiani, S.G. Karadogan, T. Andersen, J. Larsen, L.K. Hansen. The Role of Top-Down Attention in the Cocktail Party: Revisiting Cherry's Experiment after Sixty Years. ICMLA, Int. Conf. on Machine Learning and Applications (2012).

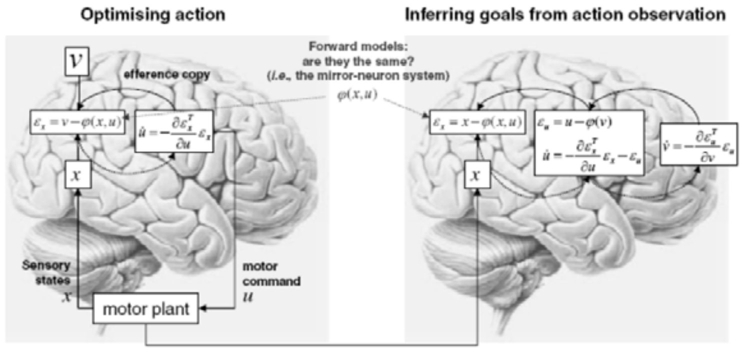
Cognitive modeling, mental models

Human cognition is often to act on *weak signals*, i.e., lack of information or poor signal to noise conditions.

Solve the problem by being very sensitive and post-process alarms with rich context models

Mental models can be more or less well-aligned with actual physics /ecology, c.f. Friston et al.'s *Predictive coding* model

J.M.Kilner, K.J.Friston C.D.Frith.
Predictive coding: an account of the mirror neuron system.
Cogn Process. 2007 Sep;8(3):159-66



Simplified scheme for **motor control**: The motor plant receives commands (u) and changes the sensory input (x). These commands are constructed by a controller (inverse model) to minimise the difference between the desired trajectory of the states (v) and those predicted by the forward model. The forward (predictor) model is a function of the [efference] copy of the motor command. In this case, the goal is known and only u is optimised. The inverse model or controller is represented as a recognition function that minimises prediction error by gradient descent (the dot above a variable means rate of change).

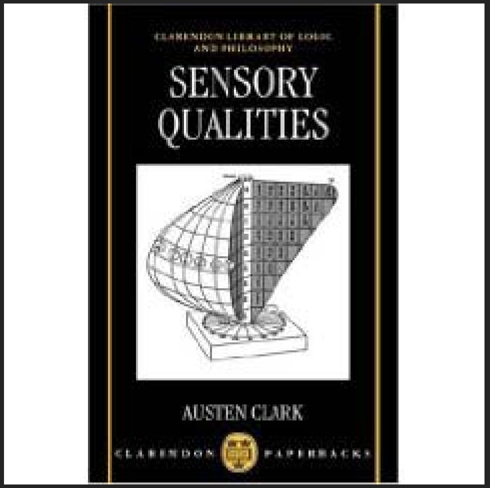
Simplified scheme for **action-perception**: A hierarchical generative or forward model of sensory states is inverted to infer their [unknown] causes. These causes include the motor commands (u) of the observed agent that are inferred by minimising the difference between the observed and predicted states (using a forward model of the motor plant). The agent's goals are inferred by minimising the error between the inferred commands (u) and those predicted by their forward model, which is a function of goals.

Subjective, sensory data

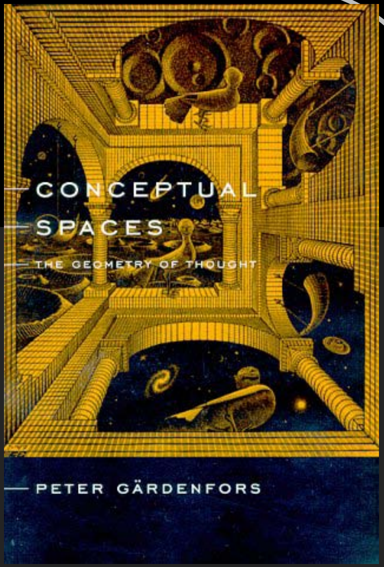
Qualitative data often mapped with MDS multidimensional scaling: low-dimensional, neighbor preserving Euclidean representation

Austen Clark in Sensory Qualities (1993):

> "*The number of dimensions of the MDS space corresponds to the number of independent ways in which stimuli in that modality can be sensed to resemble or differ, but the dimensions per se have no meaning*"



Gärdenfors' conceptual spaces



Cognitive models:

– Symbolic, associative/connectionist, geometrical

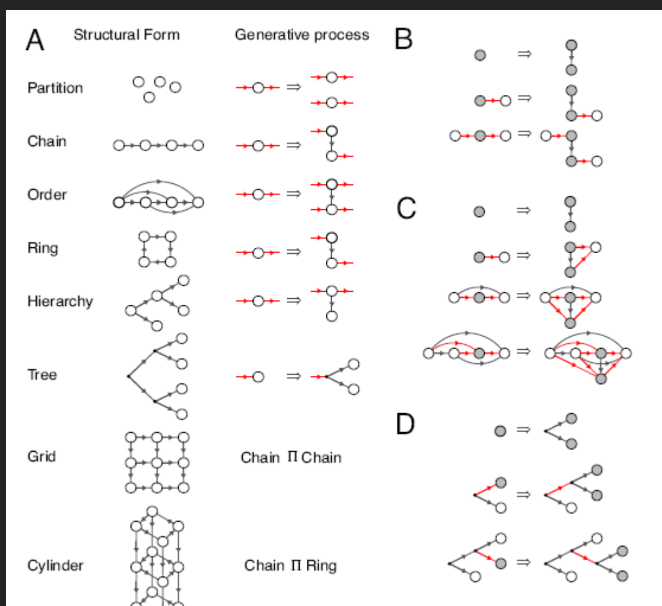
Human cognition ~ similarity judgments ~ Gestalt theory ~ geometrical proximity

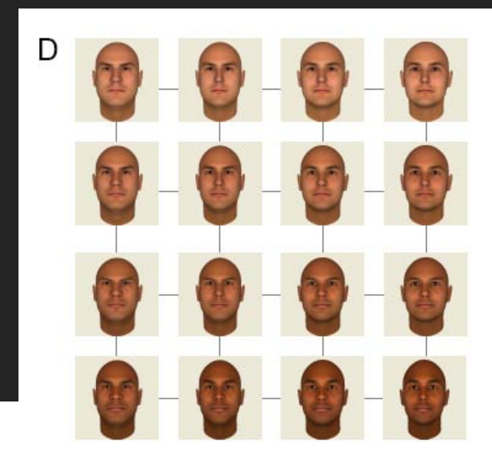
How to identify conceptual spaces, i.e., geometrical representations? - Cognitive component analysis?

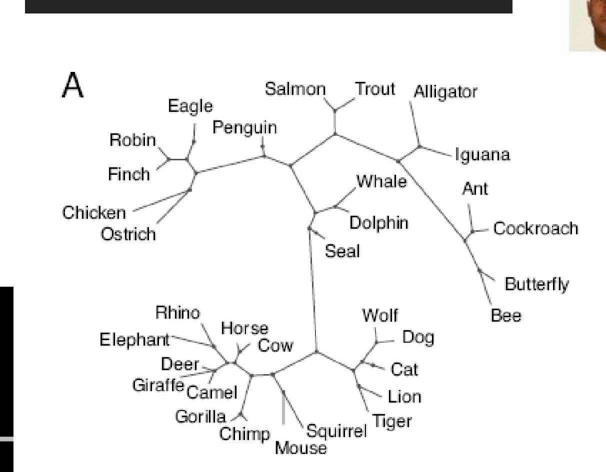
(Gärdenfors, 2000)

Kemp-Tenenbaum – Discovery of structural form
(2008)

Human mind has access only to relatively low complexity modeling tools

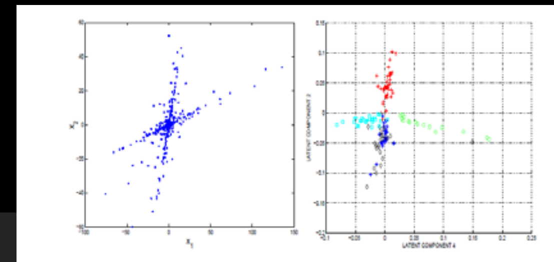






DTU

Cognitive compatibility



Unsupervised Learning

Hidden variable

$$p(\mathbf{s} \mid \mathbf{x}, \mathbf{w}_u) \propto p(\mathbf{x} \mid \mathbf{s}, \mathbf{w}_u) p(\mathbf{s} \mid \mathbf{w}_u)$$

”Cognitive event”:
Data, sound, image, behavior

Supervised learning

$$p(\mathbf{y} \mid \mathbf{x}, \mathbf{w}_s)$$

”Cognitive” label, i.e. provided by a human observer

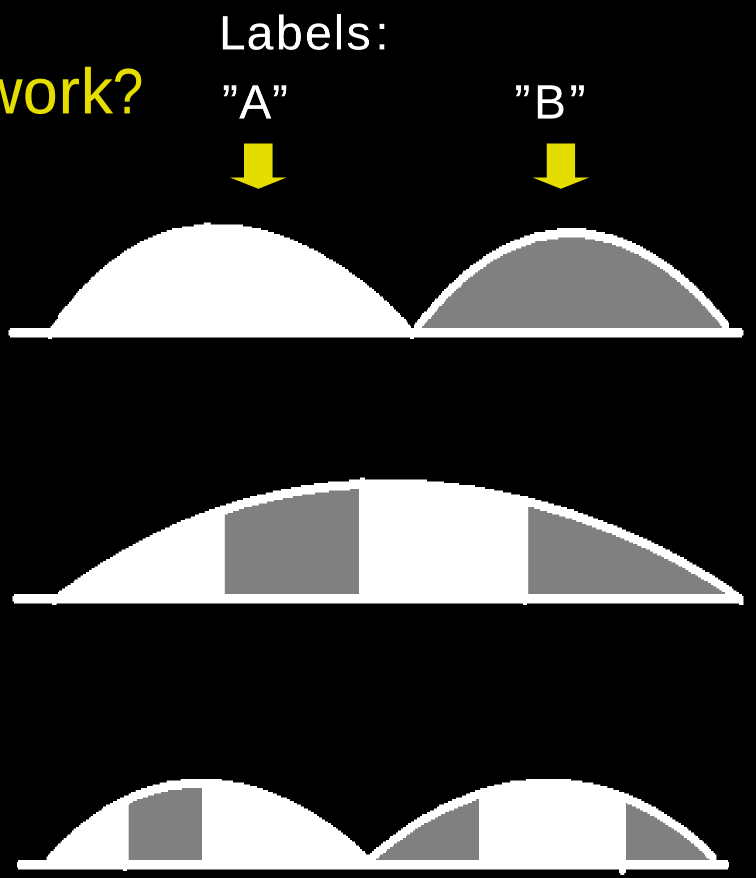
How well do these learned representations match: s = y ?

When can COCA be expected to work?

If "statistical structure" in the relevant feature space is well aligned with the label structure we expect high **cognitive compatibility**

Unsupervised-then-supervised learning can explain "learning from a single example"

The Good,
the Bad,
and the Ugly...

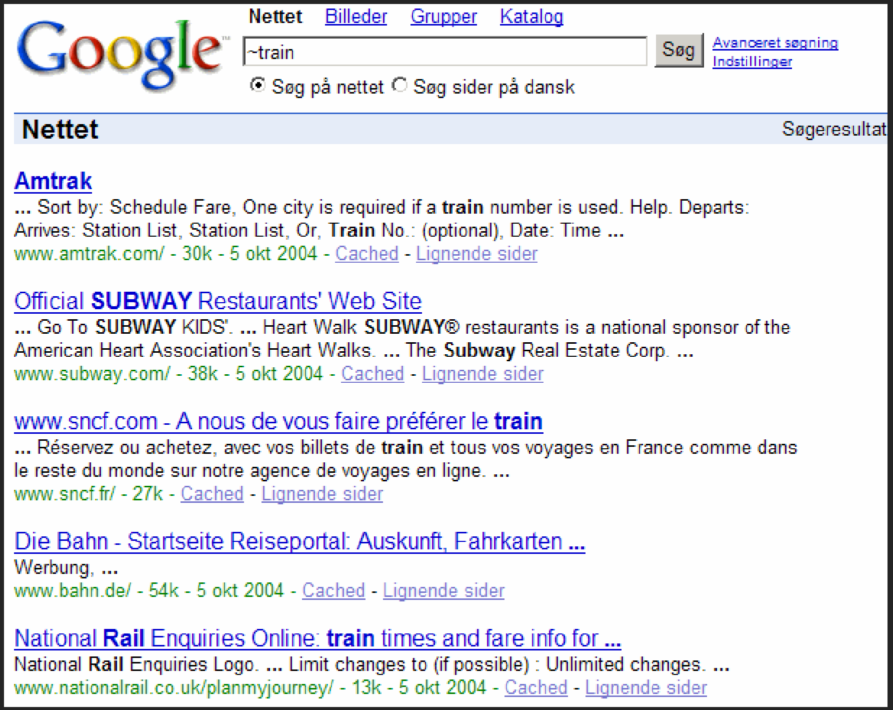


How will COCA help computers understand media content?

Understand = simulate cognitive processing in humans

Help metadata estimation automatic tagging of digital media (sound/images/video/ deep web objects)

Basic signal processing tools are known (perceptual models...)



Vector space representation

Abstract representation - can be used for all digital media

A “cognitive event” is represented as a point in a high-dimensional ”feature space” – document similarity ~ spatial proximity in a given metric

Text: Term/keyword histogram, N-grams

Image: Color histogram, texture measures

Video: Object coordinates (tracking), active appearance models

Sound: Spectral coefficients, mel cepstral coefficients, gamma tone filters

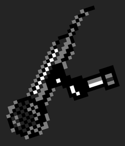
Contexts can be identified by their feature associations (= Latent semantics)

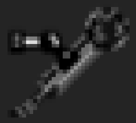
S. Deerwester et al. *Indexing by latent semantic analysis*.
Journal of the American Society for Information Science, 41(6), 391-407, (1990)

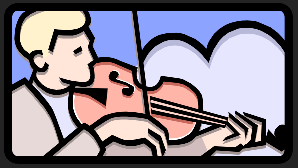
The independent context hypothesis: The perpetual cocktail party

Challenge: Presence of multiple agents/contexts
Need to ”blindly” separate source signals = learn contexts
Independent Component Aanalysis come to rescue!

$$x(feature, time) = \sum_k A(feature, k)\ s(k, time)$$







Linear mixing generative model ICA - “Synthesis”
simplistic model incorporating sparsity and independence

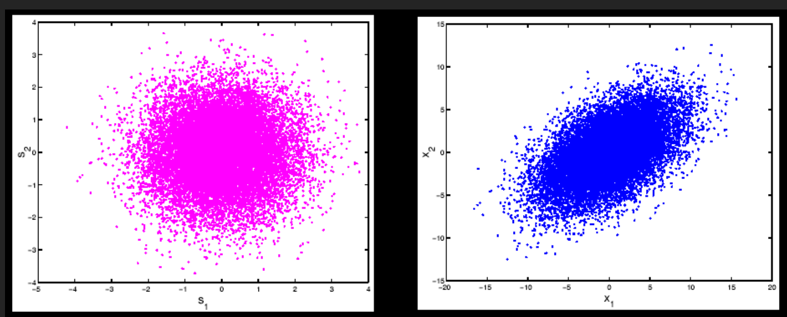
Space-time matrix

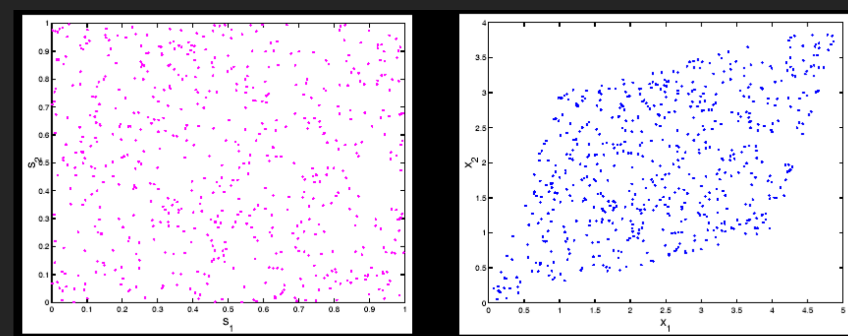
Component’s “where”

Vector of “what”

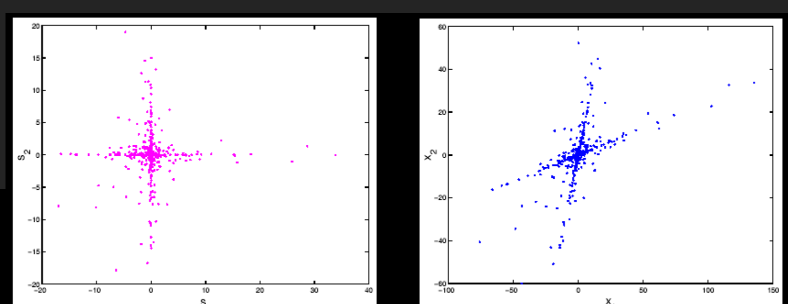
$$x(loc,time) = \sum_k A(loc,k)\ s(k,time)$$

Normal sources





Dense sources



Sparse sources

Protocol for comparing supervised and unsupervised learning

Use an “unsupervised-then-supervised” classifier:

- Train the unsupervised scheme, eg., ICA
- Freeze the ICA representation (A matrix)
- Train a simple (e.g. Naïve Bayes) classifier using the features obtained in unsupervised learning Use

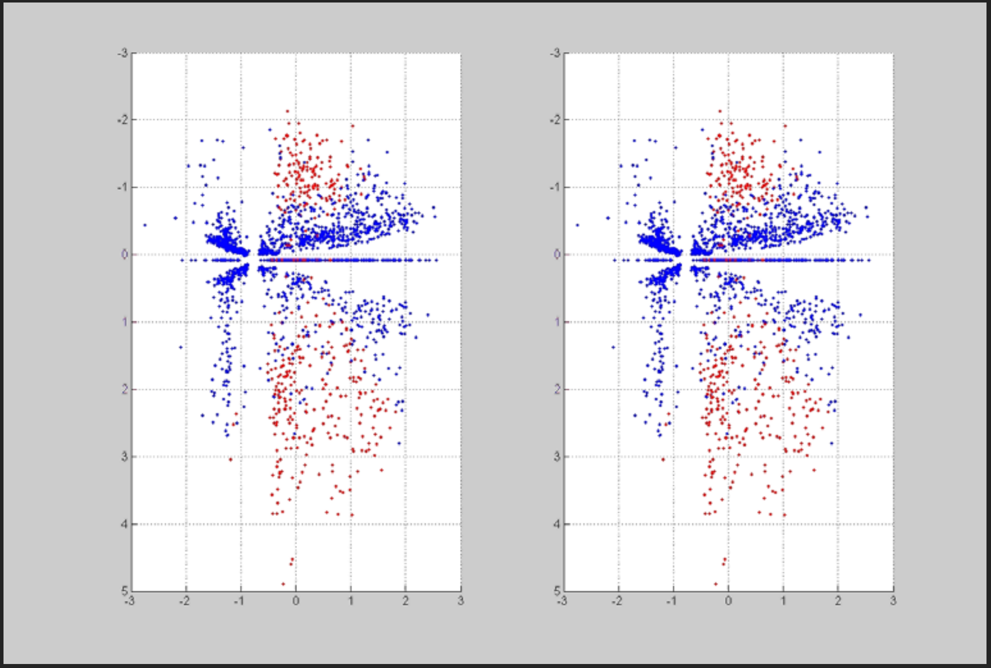
Compare with supervised classifier == human proxy

- Error rates of the two systems
- Compare posterior probabilities

Research question: Can statistics of independence account for human object detection/uncertainty?

Phoneme classification

Nasal vs oral: ”Esprit project ROARS” (Alinat et al., 1993)



Binary classification

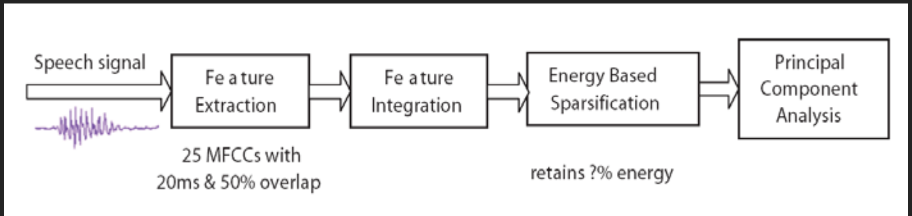
Error rates: **0.23** (sup.), **0.22** (unsup.)

Bitrates: **0.48** (sup.), **0.39** (unsup.)

Cognitive components of speech

Basic representation: Mel weigthed cepstral coefficients (MFCCs)

Modeling at different time scales 20 msec – 1000 msec



Phonemes

Gender

Speaker identity

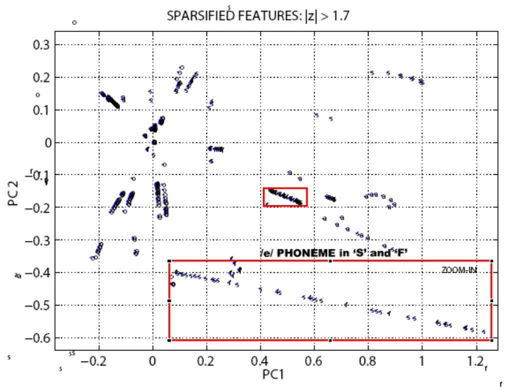
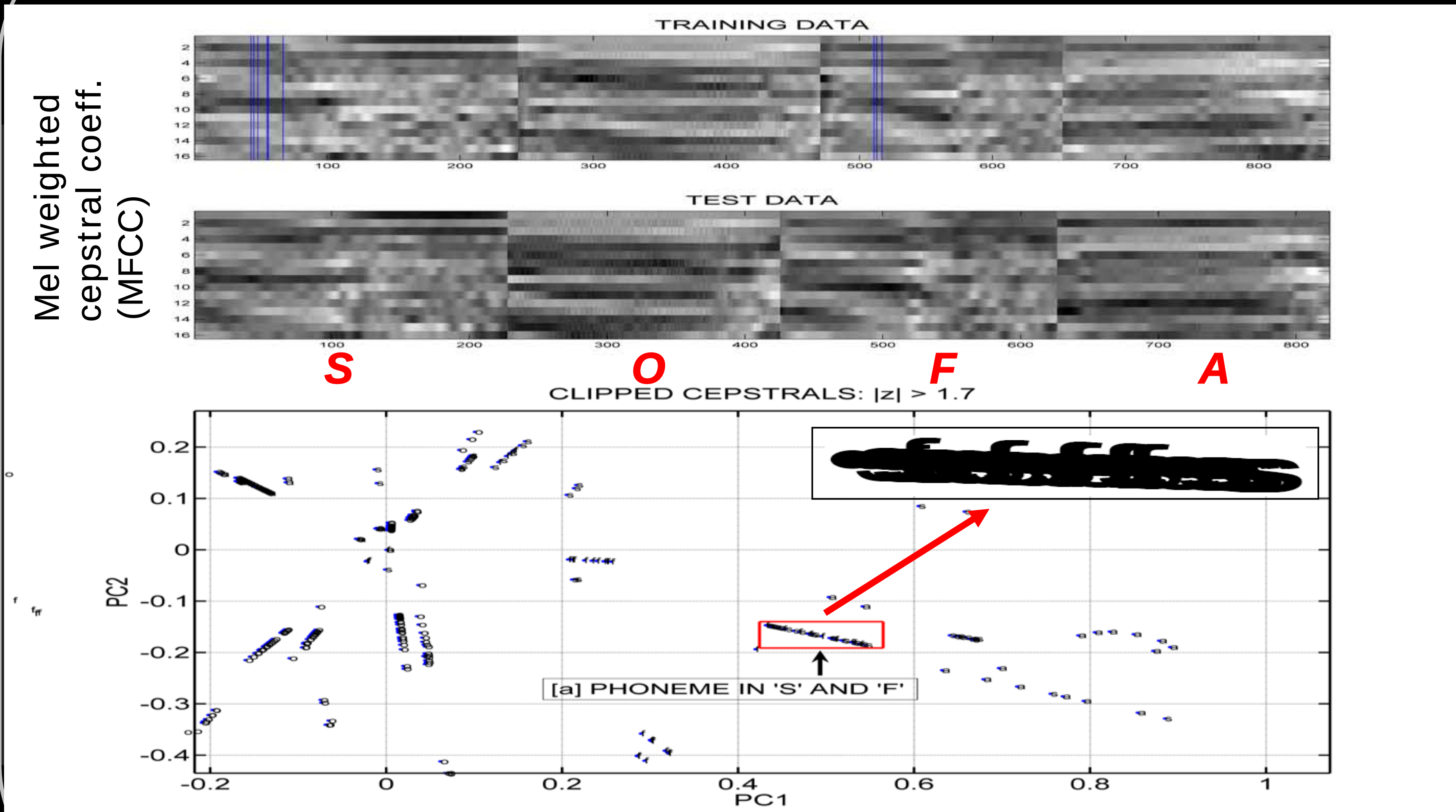


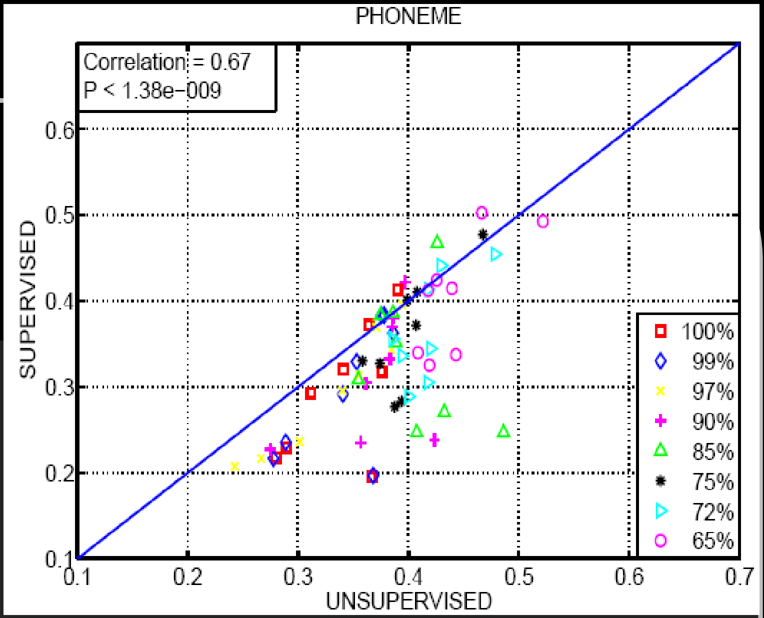
Figure 3: The latent space is formed by the two first principal components of data consisting of four separate utterances representing the sounds 's', 'o', 'f', 'a'. The structure clearly shows the sparse component mixture, with 'rays' emanating from the origin (0,0). The ray embraced in a rectangle contains a mixture of 's' and 'f' features, a cognitive component associated with the vowel /e/ sound.

Co-worker: Ling Feng


Mel weighted cepstral coeff. (MFCC)
TRAINING DATA
TEST DATA
S
O
F
A
CLIPPED CEPSTRALS: |z| > 1.7
PC2
PC1
[a] PHONEME IN 'S' AND 'F'

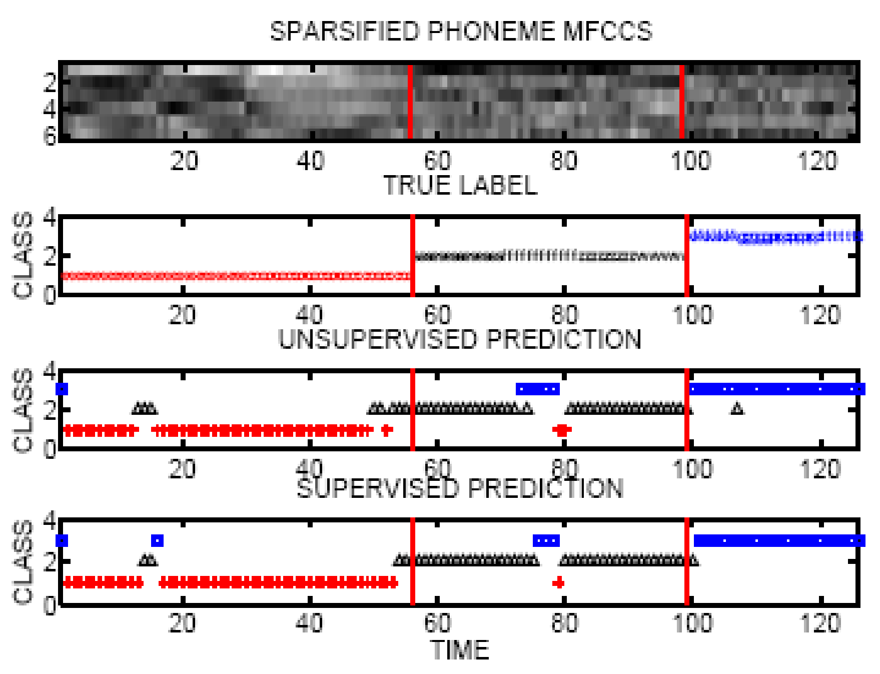
Error rate comparison

For the given time scales and thresholds, data locate around $y = x$, and the correlation coefficient $\rho=0.67$, $p<1e-09$.

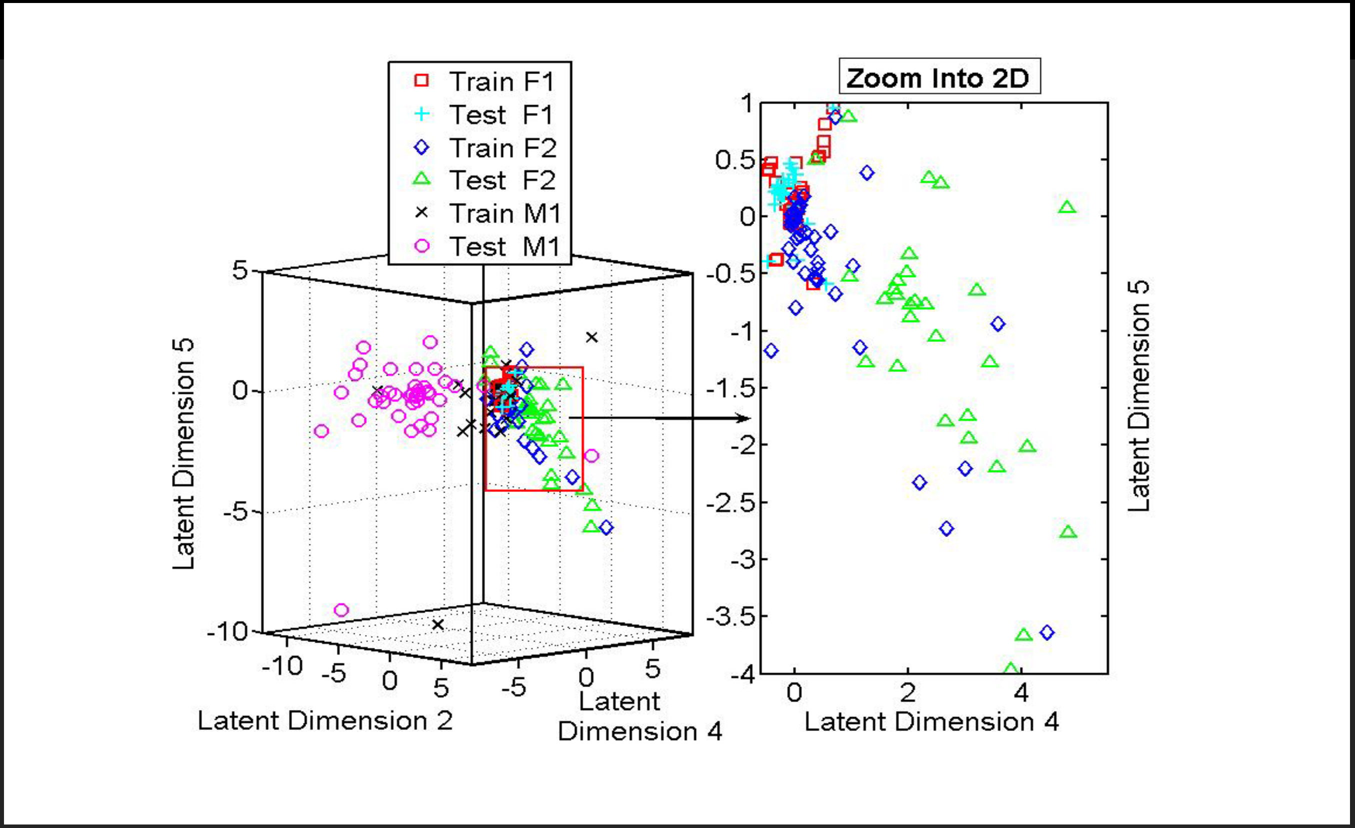


Sample-to-sample correlation

- Three groups: vowels eh, ow; fricatives s, z, f, v; and stops k, g, p, t.
- 25-d MFCCs; EBS to keep 99% energy; PCA reduces dimension to 6.
- Two models had a similar pattern of making correct predictions and mistakes. Match between supervised and unsupervised learning = 91%.



Longer time scales



Time integrated (1000ms) MFCC's: text independent speaker recognition….

Feng & Hansen (CIMCA, 2005)

Error rate correlations
for super/unsupervised
learning for different cognitive
time scales and events

Challenged by degree of
sparsity and time averaging

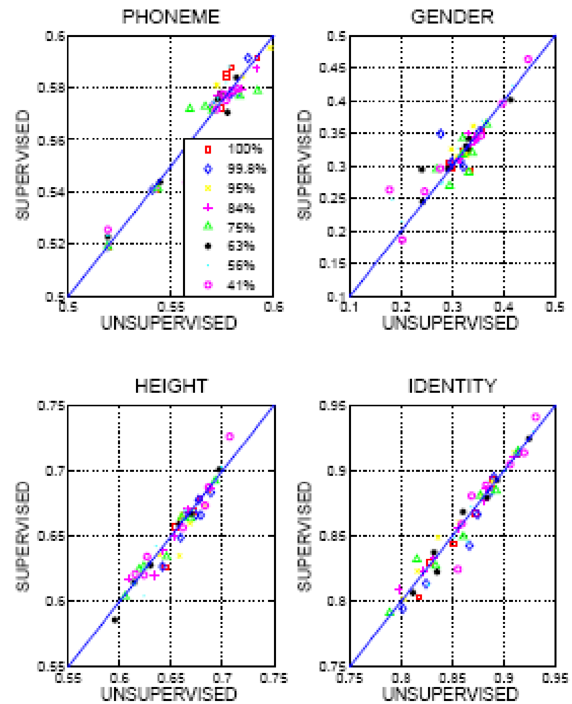


Fig. 4. Figure shows test error rates of both supervised and unsupervised learning on four topics: phonemes, gender, height and identity. Solid lines indicate $y = x$ in the coordinate systems. All data located along this line, meaning high correlation between supervised and unsupervised learning.

"Higher" cognitive representations: Digital media vector space representation

Abstract representation - can be used for all digital media

Document is represented as a point in a high-dimensional "feature space"
document similarity ~ spatial proximity in a given metric

Text: Term/keyword histogram, N-grams
Image: Color histogram, texture measures
Video: Object coordinates (tracking), active appearance models
Sound: Spectral coefficients, cepstral coefficients, gamma tone filters

Latent semantics

Document features are correlated, the pattern of correlation reflects "associations".

Associations are context specific

Word sets are activated in concert in a given context

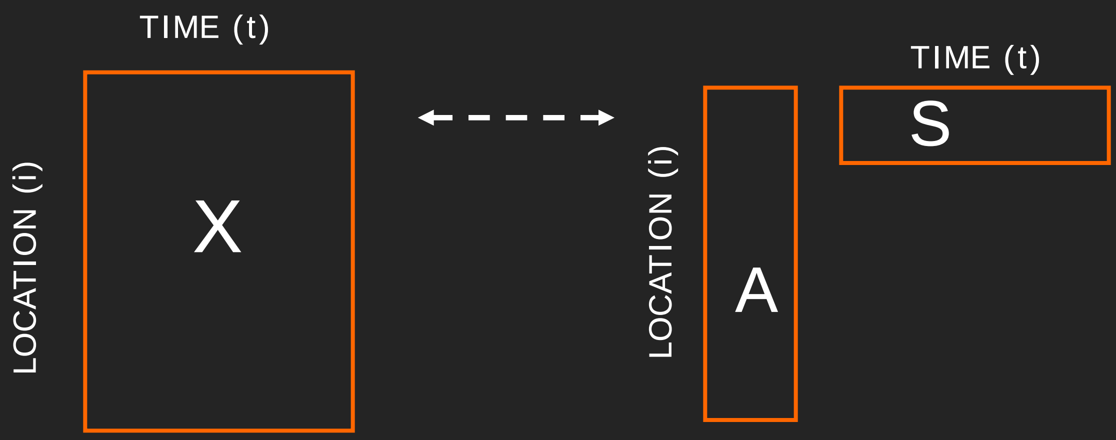
ape ~ zoo, zoo ~ elephant => ape ~ elephant

Latent semantic analysis: Contexts can be identified by term co-variance patterns (PCA)

Deerwester, S., Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R:
Indexing by latent semantic analysis.
Journal of the American Society for Information Science, 41(6), 391-407, (1990)

Factor models for what / where

Represent a datamatrix by a low-dimensional approximation



Generative model for hidden variables

$$\mathbf{x} = \mathbf{A}\mathbf{s} + \boldsymbol{\varepsilon}, \quad \boldsymbol{\varepsilon} \sim N(\mathbf{0}, \boldsymbol{\Sigma})$$

$$p(\mathbf{x} \mid \mathbf{A}, \boldsymbol{\theta}, \boldsymbol{\Sigma}) = \int p(\mathbf{x} \mid \mathbf{A}, \mathbf{s}, \boldsymbol{\Sigma}) \boxed{p(\mathbf{s} \mid \boldsymbol{\theta})} d\mathbf{s}$$

$$p(\mathbf{x} \mid \mathbf{A}, \mathbf{s}, \boldsymbol{\Sigma}) = \left| 2\pi\boldsymbol{\Sigma} \right|^{-1/2} e^{-\frac{1}{2}(\mathbf{x}-\mathbf{A}\mathbf{s})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x}-\mathbf{A}\mathbf{s})}$$

Source distribution:
PCA: ... normal
ICA: ... other
IFA: ... Gauss. Mixt.

PCA: $\boldsymbol{\Sigma} = \sigma^2 \cdot \mathbf{1},$

FA: $\boldsymbol{\Sigma} = \mathbf{D}$

S known:	GLM
(1-A^{-1}) sparse:	SEM
S,A positive:	NMF

Højen-Sørensen, Winther, Hansen,
Neural Comp (2002), Neurocomputing (2002)

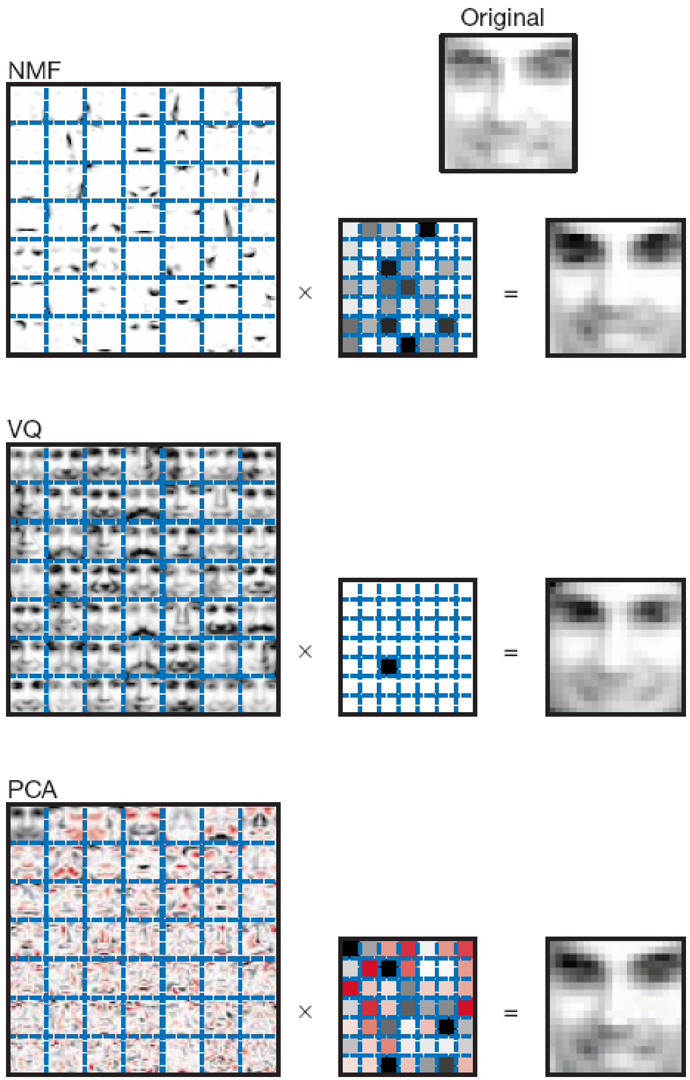


Figure 1 Non-negative matrix factorization (NMF) learns a parts-based representation of faces, whereas vector quantization (VQ) and principal components analysis (PCA) learn holistic representations. The three learning methods were applied to a database of $m = 2{,}429$ facial images, each consisting of $n = 19 \times 19$ pixels, and constituting an $n \times m$ matrix V. All three find approximate factorizations of the form $V \approx WH$, but with three different types of constraints on W and H, as described more fully in the main text and methods. As shown in the 7×7 montages, each method has learned a set of $r = 49$ basis images. Positive values are illustrated with black pixels and negative values with red pixels. A particular instance of a face, shown at top right, is approximately represented by a linear superposition of basis images. The coefficients of the linear superposition are shown next to each montage, in a 7×7 grid, and the resulting superpositions are shown on the other side of the equality sign. Unlike VQ and PCA, NMF learns to represent faces with a set of basis images resembling parts of faces.

Learning the parts of objects by non-negative matrix factorization

Daniel D. Lee* **& H. Sebastian Seung***†

* *Bell Laboratories, Lucent Technologies, Murray Hill, New Jersey 07974, USA*
† *Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

NATURE | VOL 401 | 21 OCTOBER 1999 | www.nature.com

Modeling the generalizability of factorization

Rich physics literature on ”retarded” learning

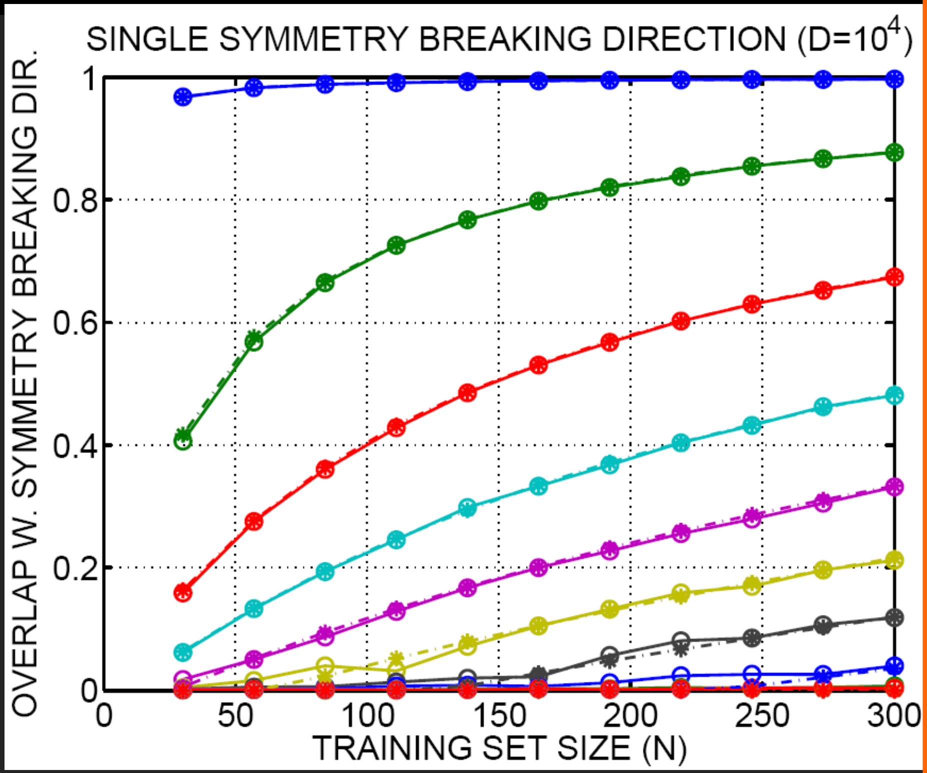
Universality

- Generalization for a ”single symmetry breaking direction” is a function of ratio of N/D and signal to noise S
- For subspace models-- a bit more complicated -- depends on the component SNR’s and eigenvalue separation
- For a single direction, the mean squared overlap $R^2 = \langle (u^T_1 * u_0)^2 \rangle$ is computed for N,D -> ∞

$$R^2 = \begin{cases} (\alpha S^2 - 1)/S(1+\alpha S) & \alpha > 1/S^2 \\ 0 & \alpha \le 1/S^2 \end{cases}$$

$$\alpha = N/D \qquad S = 1/\sigma^2 \qquad N_c = D/S^2$$

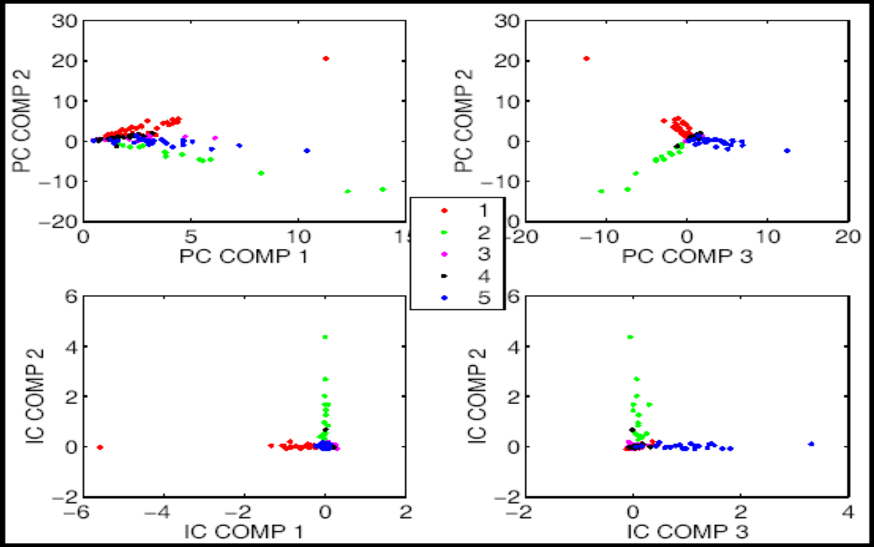
Hoyle, Rattray: Phys Rev E 75 016101 (2007)

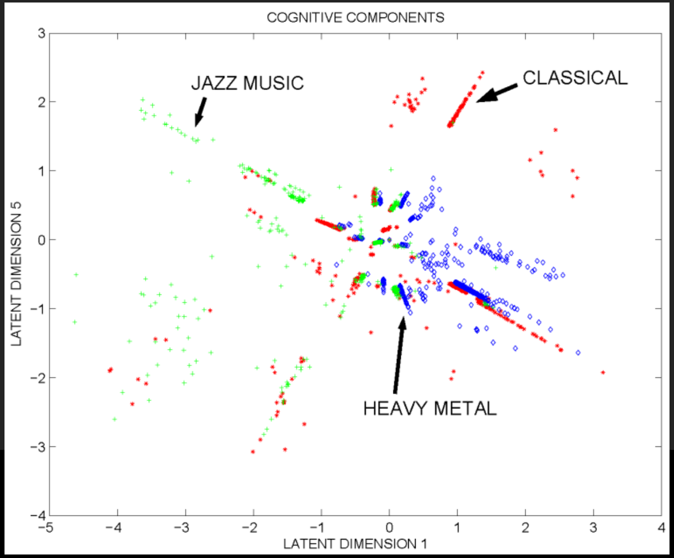


N_c = (0.0001, 0.2, 2, 9, 27, 64, 128, 234, 400, 625)

σ = (0.01, 0.06, 0.12, 0.17, 0.23, 0.28, 0.34, 0.39, 0.45, 0.5)

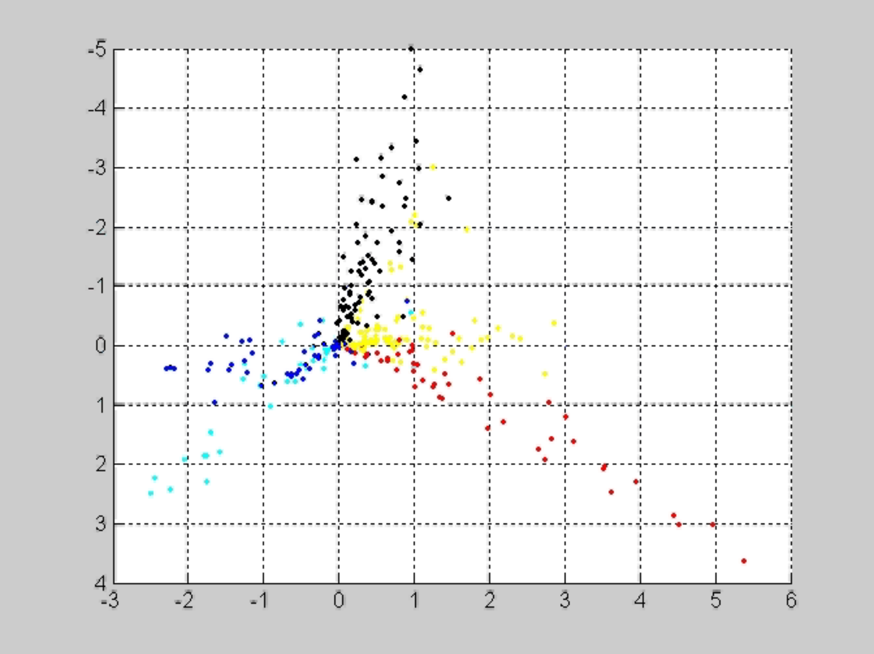
Linear mixture of independent agents in term-document scatterplots





Linear mixture of independent contexts observed in short time features (mel-ceptrum) in a music database.

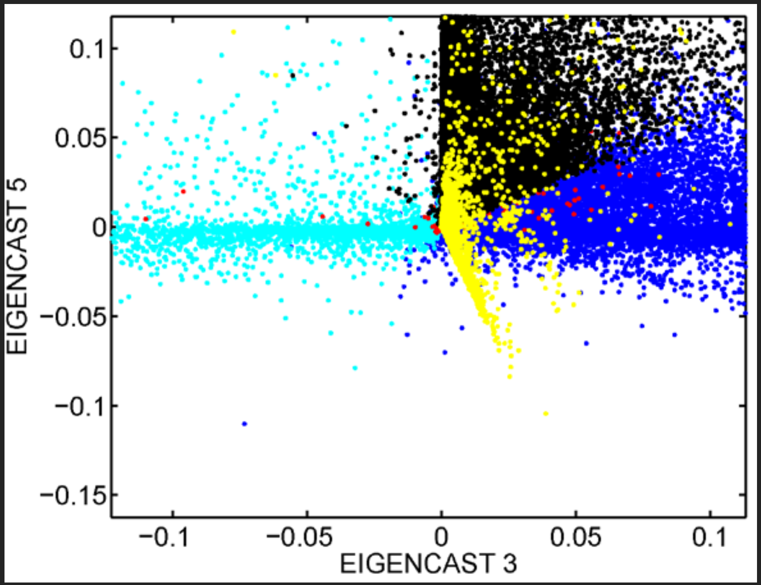
Complex (social) networks: Linear mixtures of independent ”interest”?



Genre patterns in expert opinion (tags) on 400 musical artists

Berenzweig,. Logan,. Ellis, Whitman (2004). A large-scale evaluation of acoustic and subjective music-similarity measures *Computer Music Journal*, 28(2), 63-76, 2004

”Movie actor network”
- A collaborative small world network



128.000 movies

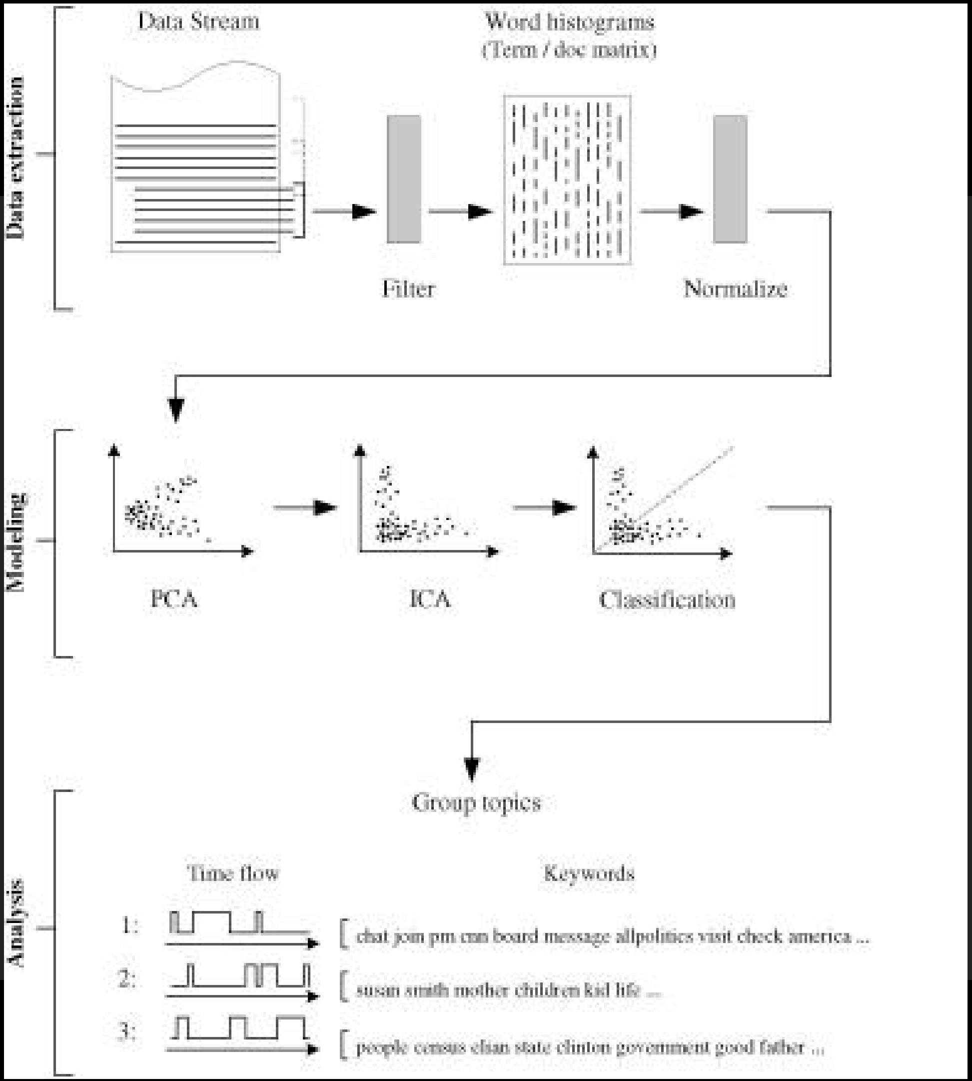
380.000 actors

Independent contexts in document databases

• x(j,t) is the occurence of the j'th word in the t'th document.

• s(k,t) quantifies how much the k'th context is expressed in t'th document.

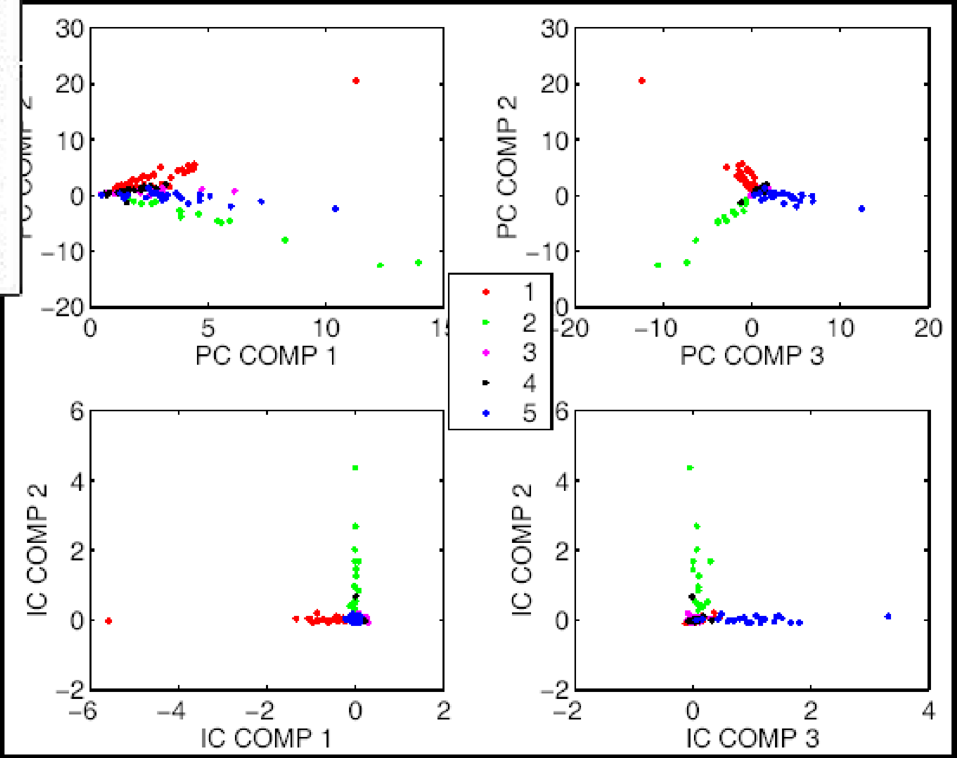
• A(j,k) quantifies the typical importance of the j'th word in the k'th context



ICA in text
Isbell and Viola (1999)
Kolenda, Hansen, (2000)

PCA vs ICA document scatterplots

Terms	Documents								
	c1	c2	c3	c4	c5	m1	m2	m3	m4
computer	1	1	0	0	0	0	0	0	0
EPS	0	0	1	1	0	0	0	0	0
human	1	0	0	1	0	0	0	0	0
interface	1	0	1	0	0	0	0	0	0
response	0	1	0	0	1	0	0	0	0
system	0	1	1	2	0	0	0	0	0
time	0	1	0	0	1	0	0	0	0
user	0	1	1	0	1	0	0	0	0
graph	0	0	0	0	0	0	1	1	1
minors	0	0	0	0	0	0	0	1	1
survey	0	1	0	0	0	0	0	0	1
trees	0	0	0	0	0	1	1	1	0



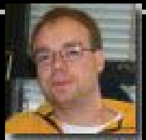
Independent contexts in dynamic text: Chat room analysis

We logged a days
chat in a CNN ”news cafe”.

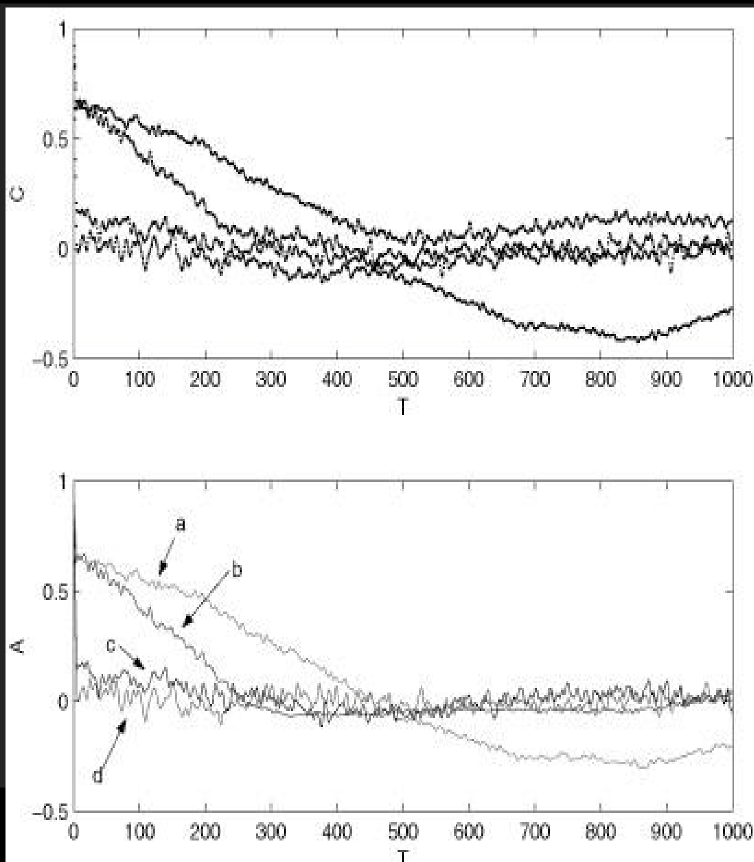
The database involves
120 users chatting
during an 8 hour period

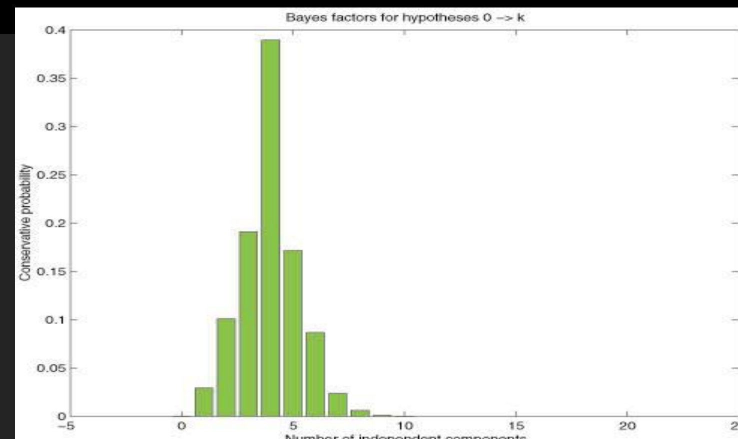
... just statements like that - over the past
few weeks.
<Miez> heyy seagate
<Recycle> denise: he deserved it for stealing os code in his early days
<Zeno> ok Sharonelle
<denise> LOL @ Recycle
<HaleyCNN> Join Book chat at 10am ET in #auditorium. Chat with Robert Ballard author of "Eternal Darkness: A Personal History of Deep-Sea Exploration," after his appearance on CNN Morning News at 9:30am ET.
<heartattackagain> Ed Shore....lol....We might have an operating system that doesn't crash every thirty minits....lololol.....
<EdShore> Shooby, I don't believe you. I've been doing this sine PET, TRS-80, and PIRATES! Don't tell me you've been CHATTING! PROVE IT!
<Zeno> Recycle LOL ethical and criminal laws are different for the business world
<_Seagate_> Recycle, thats what the technology business is all about.
<tribe> I heard a local radio talk show host saying last night that he has noticed everytime this Elian issue slows down, something happens to either the family in Miami or in Cuba to put it right back in the headlines. He mentioned the cousin's hospitalization as just the latest saga
<Diogenes> If Bill Gates was in Silicon Valley never a word would you have ever heard.
<Zeno> EdC you may have been doing sine but i have been doing cosine.
<shooby> EdShore: Compuserve since, heck, 76?
<Zeno> i mean EdShore
<Recycle> rumor has it that he was even dumpster diving

Coworkers: Thomas Kolenda,
Jan Larsen




ICA by dynamic decorrelation





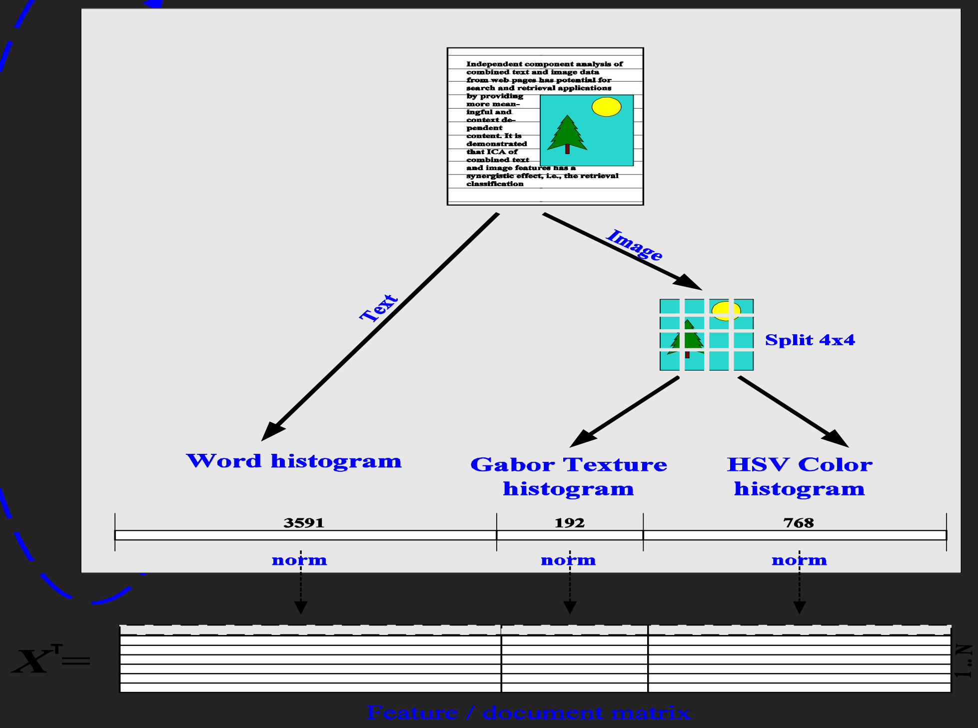
The Bayes factor - P(M|D) - of each model is estimated in the BIC approximation

Source autocorrelations

Independent contexts in multi-media

- Organizing webpages in categories
- Labels obtained from
  Yahoo’s directory

- Features: Text, color, and texture subsets of MPEG image features



L.K. Hansen, J. Larsen and T. Kolenda “On Independent Component Analysis for Multimedia Signals”.
In L. Guan et al.: *Multimedia Image and Video Processing*, CRC Press, Ch. 7, pp. 175-199, 2000.

Coworkers: Thomas Kolenda,
Jan Larsen

Performance of the system trained by associating unsupervised independent components with labels – generalization based on Yahoo cathegories

Modality	Classification Error
Color	23.0%
Texture	18.0%
Texture/Color	11.5%
Text	5.7%
Combined (texture/color/text)	2.8%

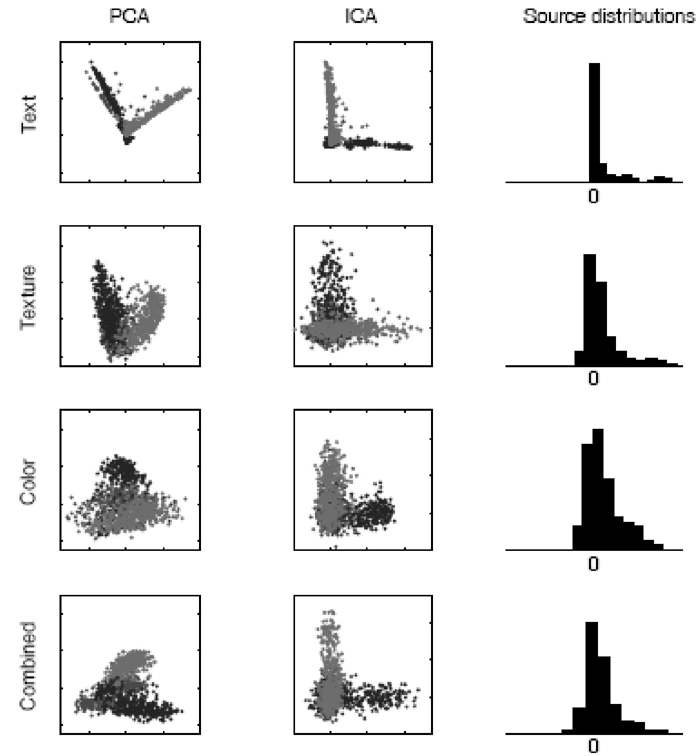


Fig. 3. Scatterplots of the text and image multimedia data, projected to a two-dimensional subspace found by PCA. Grey value of points corresponds to the three classes considered, see Fig. 4. The ray like structure strongly suggest an ICA interpretation, however, the relevance of this representation can only be determined by a subsequent inspection of the recovered source signals. As we will see in section 4.6, it turns out that there is an interesting alignment of the source signals and a manual labeling of the multimedia documents.

Texture (K=13)

69.75	7.75	6.5
11.5	88.5	5.75
18.75	3.75	87.75

Text (K=45)

93	2	2.25
0.5	94.75	2.5
6.5	3.25	95.25

Texture Color (K=26)

82	1.75	4.5
9	93.75	5.75
9	4.5	89.75

Texture (K=13)

69.75	7.75	6.5
11.5	88.5	5.75
18.75	3.75	87.75

Color (K=16)

70.75	3.75	10
12	81.5	11.25
17.25	14.75	78.75

Text (K=45)

93	2	2.25
0.5	94.75	2.5
6.5	3.25	95.25

Combined errorrate: 2.8%

Single best errorrate: 5.7%

Texture Color (K=26)

82	1.75	4.5
9	93.75	5.75
9	4.5	89.75

Texture Color Text (K=26)

98.25	0.25	0.75
0.75	98	3.75
1	1.75	95.5

ht guard

CASTSEARCH - CONTEXT BASED SPEECH DOCUMENT RETRIEVAL

Lasse Lohilahti Mølgaard, Kasper Winther Jørgensen, and Lars Kai Hansen

Informatics and Mathematical Modelling
Technical University of Denmark Richard Petersens Plads
Building 321, DK-2800 Kongens Lyngby, Denmark

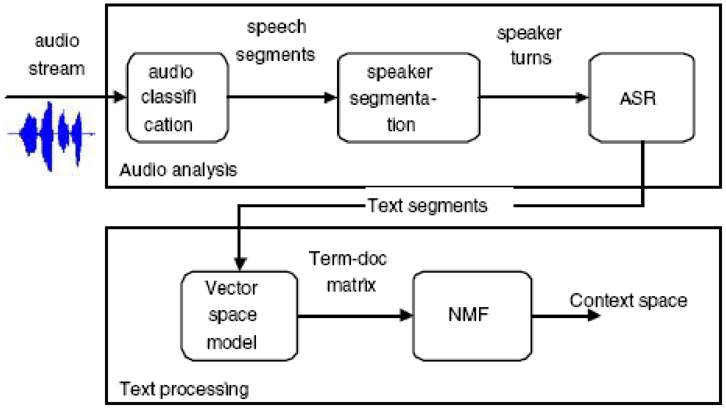
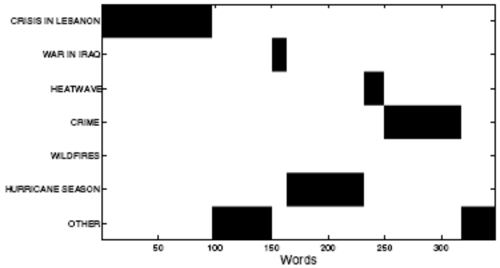
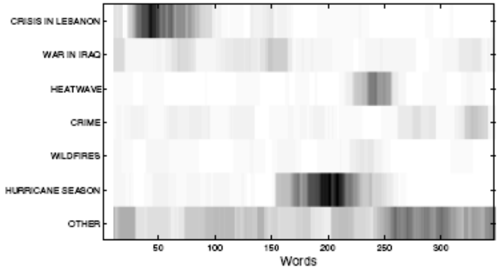


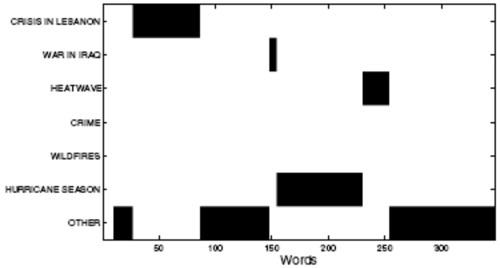
Fig. 1. The system setup. The audio stream is first processed using audio segmentation. Segments are then using an automatic speech recognition (ASR) system to produce text segments. The text is then processed using a vector representation of text and apply non-negative matrix factorization (NMF) to find a topic space.



(a) Manual segmentation.



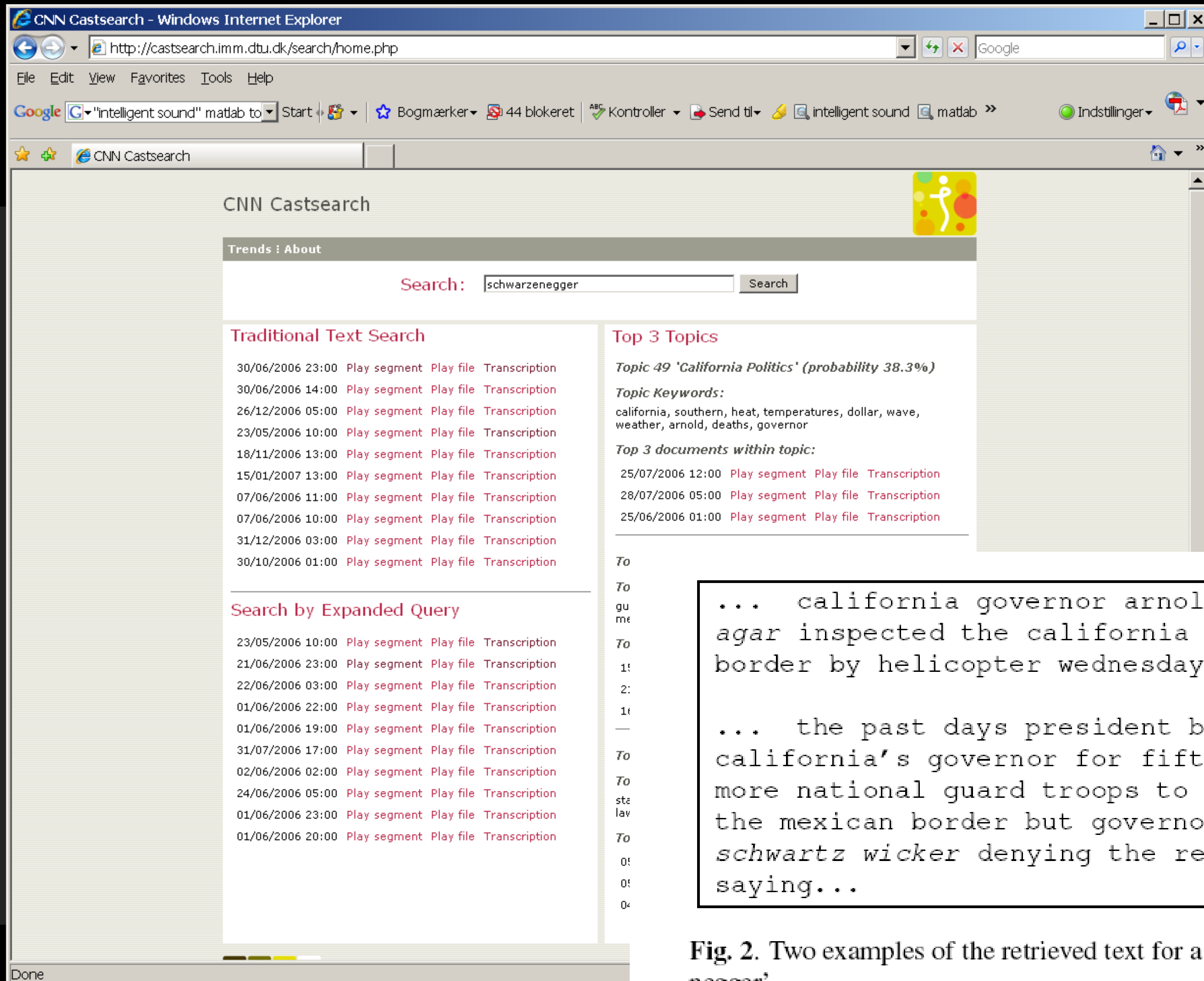
(b) $p(k|d^*)$ for each context. Black means high probability.



(c) The segmentation based on $p(k|d^*)$.

Fig. 3. Figure 3(a) shows the manual segmentation of the news show into 7 classes. Figure 3(b) shows the distribution $p(k|d^*)$ used to do the actual segmentation shown in figure 3(c). The NMF-segmentation is in general consistent with the manual segmentation. Though, the segment that is manually segmented as 'crime' is labeled 'other' by the NMF-segmentation

Mølgaard et al. 2007



... california governor arnold's *fortson agar* inspected the california mexico border by helicopter wednesday to see ...

... the past days president bush asking california's governor for fifteen hundred more national guard troops to help patrol the mexican border but governor orville *schwartz wicker* denying the request saying...

Fig. 2. Two examples of the retrieved text for a query on 'schwarzenegger'.

castsearch.imm.dtu.dk

MiRocket

Time frequency analysis pipeline:
MFCC's @ 30 ms windows

Temporal integration and genre classification at 1000ms

Music recommendation in 12-D genre space

•A. Meng, P. Ahrendt, J. Larsen, L.K. Hansen: *Temporal Feature Integration for Music Genre Classification. IEEE* Transactions on Audio and Speech and Language Processing 15(5): 1654-1664 (2007)

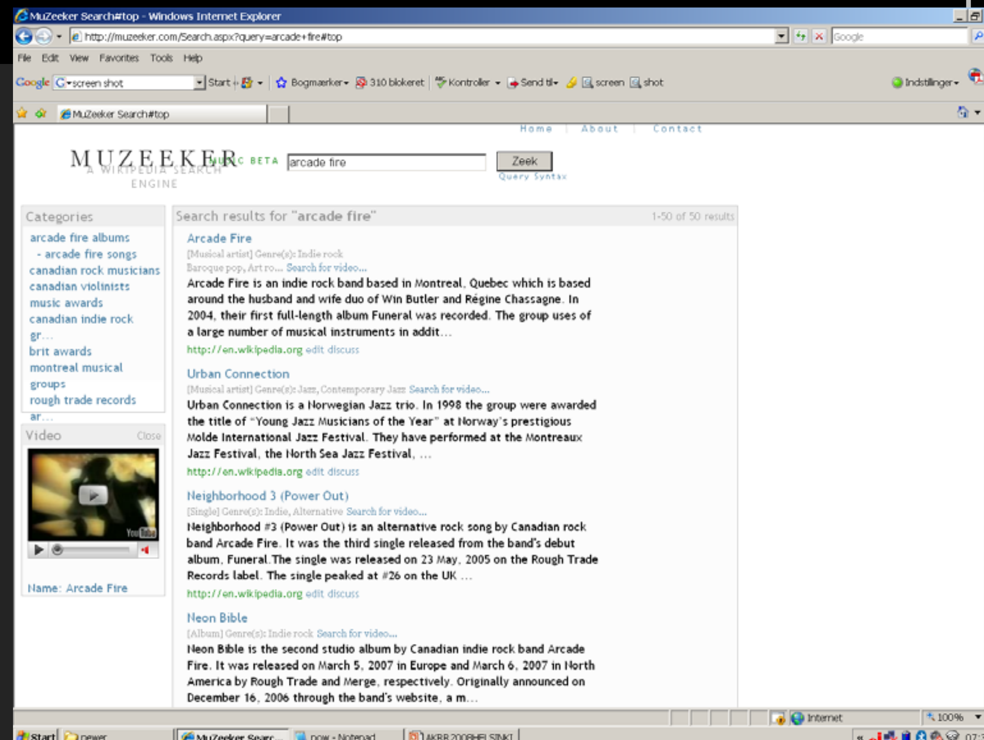
•T. Lehn-Schiøler, J. Arenas-García, K.B. Petersen and L.K. Hansen: *A Genre Classification Plug-in for Data Collection. Proc. 7th Intl. Conf. on Music Information Retrieval, ISMIR 2006, pp. 320-321, Victoria, Canada, Oct.* (2006).

•L.K. Hansen, T. Lehn-Schiøler, K.B. Petersen, J. Arenas-Garcia, J. Larsen, and S.H. Jensen: *Learning and clean-up in a large music database. EUSIPCO 2007, European Conference on Signal Processing, Poznan (2007).*

muzeeker

- Wikipedia based common sense
- Wikipedia used as a proxy for the music users mental model
- Implementation: Filter retrieval using Wikipedia's article/ categories

muzeeker.com



S. Halling, M.K. Sigurdsson, J.E. Larsen, S. Knudsen, L.K. Hansen: *MuZeeker: A domain SpecificWikipedia-based Search Engine. In Proc. First International Workshop on Mobile Multimedia Processing. Tampa, USA (2008).*

J.E. Larsen, S. Halling, M. Sigurdsson and L.K. Hansen: *MuZeeker - Adapting a music search engine for mobile phones. To appear in Springer Lecture Notes in Computer Science 'Mobile Multimedia Processing: Fundamentals,* Methods, and Applications', Selected papers from First International Workshop on Mobile Multimedia Processing, Tampa, USA. (2010).

Outlook – the future of mind reading

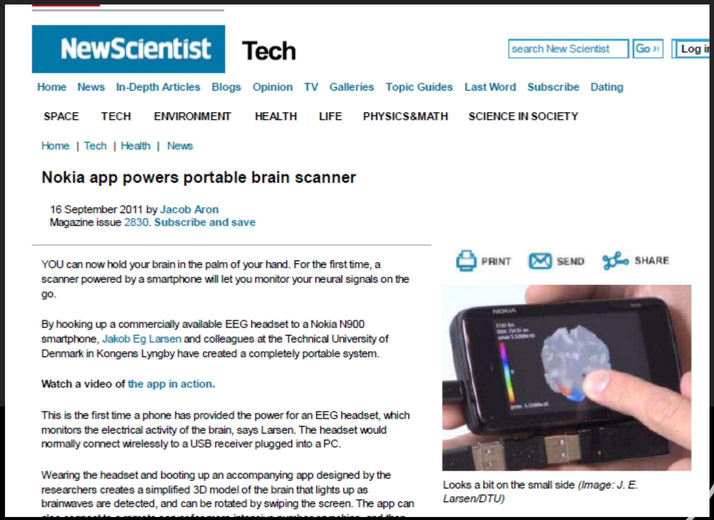
Continuous mental state
decoding in the wild

24/7 monitoring

EEG real time 3D imaging for bio-feedback

Fig. 1. Handheld brain scanner components. Emotiv EPOC wireless EEG headset (1), Emotiv Receiver module with USB connector (2), USB connector and adapter (3+4), and Nokia N900 mobile phone. The total cost of the system is less than USD1000.

hyposafe®



NewScientist Tech

search New Scientist | Go » | Log in

Home News In-Depth Articles Blogs Opinion TV Galleries Topic Guides Last Word Subscribe Dating

SPACE TECH ENVIRONMENT HEALTH LIFE PHYSICS&MATH SCIENCE IN SOCIETY

Home | Tech | Health | News

Nokia app powers portable brain scanner

16 September 2011 by Jacob Aron
Magazine issue 2830. **Subscribe and save**

PRINT SEND SHARE

YOU can now hold your brain in the palm of your hand. For the first time, a scanner powered by a smartphone will let you monitor your neural signals on the go.

By hooking up a commercially available EEG headset to a Nokia N900 smartphone, Jakob Eg Larsen and colleagues at the Technical University of Denmark in Kongens Lyngby have created a completely portable system.

Watch a video of the app in action.

This is the first time a phone has provided the power for an EEG headset, which monitors the electrical activity of the brain, says Larsen. The headset would normally connect wirelessly to a USB receiver plugged into a PC.

Wearing the headset and booting up an accompanying app designed by the researchers creates a simplified 3D model of the brain that lights up as brainwaves are detected, and can be rotated by swiping the screen. The app can

Looks a bit on the small side *(Image: J. E. Larsen/DTU)*

Quantifying subjectivity

"Affective computing"

Affective computing is research in systems that can recognize, interpret, process, and simulate human emotion

Emotions are omnipresent and extremely important to communication / opinion formation / intent reading etc

Psychology of emotion is well developed

but still far from complete...

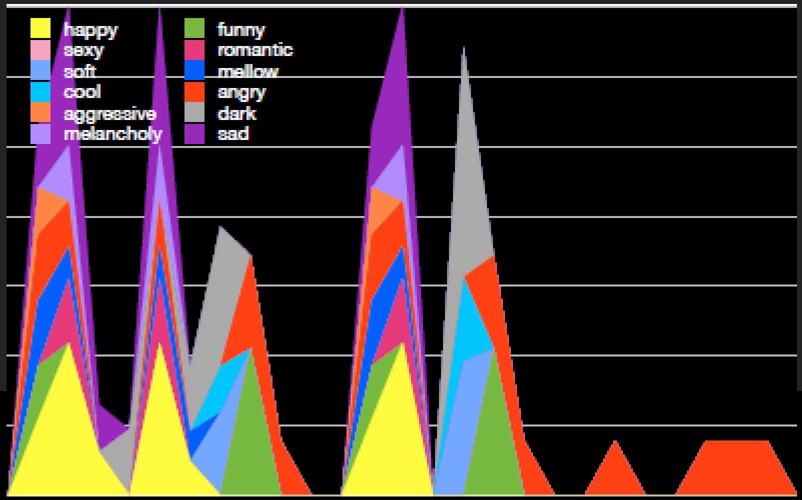
7. The law of hedonic asymmetry: *"Negative emotions last longer, positive emotions fade"*

Sentiment detection

A step towards understanding subjectivity, opinion

Important to many services
Recommender (Amazon reviews)
Information navigation, e.g. navigating music

Oasis ”Wonderwall”
Emotional content in the song lyrics through time



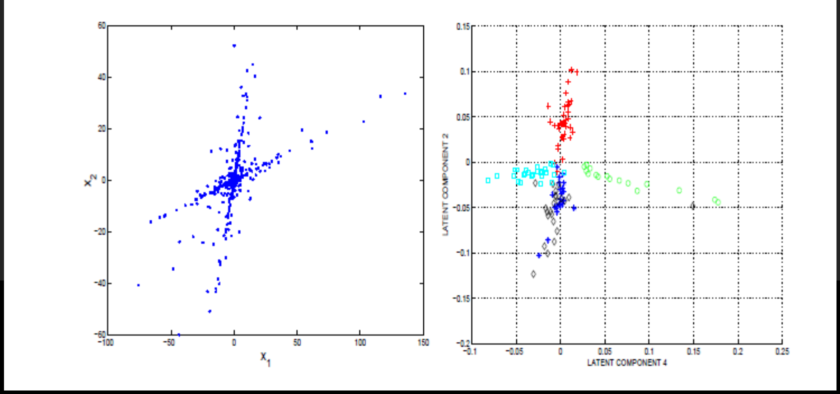
•M.K. Petersen: Modeling media as latent semantics
•based on cognitive components (Ph.D. Thesis, DTU,2010)

Conclusions & outlook

Evidence that phonemes, gender, identity are independent components 'objects' in the (time stacked) MFCC representation

Evidence that human categorization is based on sparse independent components in social networks, text, digital media

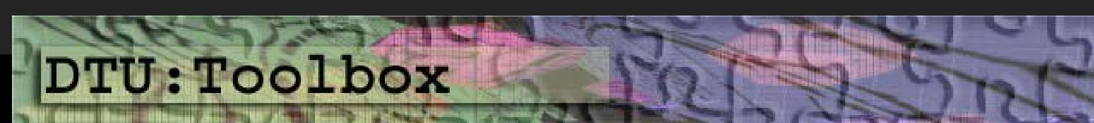
Conjecture: Objects in digital media can be identified as independent components: The brain uses old tricks from perception to solve complex "modern" problems.



Acknowledgments

- Danish Re
- EU Com
- NIH Hum

For sof
DTU:IC



Additional references

A. J. Bell and T. J. Sejnowski, “The ‘independent components’ of natural scenes are edge filters,” *Vision Research*, vol. 37, pp.3327–3338, 1997.

P. Hoyer and A. Hyvrinen, “Independent component analysis applied to feature extraction from colour and stereo images,” *Network: Comput. Neural Syst.*, vol. 11, pp. 191–210, 2000.

M. S. Lewicki, “Efficient coding of natural sounds,” *Nature Neuroscience*, vol. 5, pp. 356–363, 2002.

E. Doi and T. Inui and T. W. Lee and T. Wachtler and T. J. Sejnowski , “ Spatiochromatic Receptive Field Properties Derived from Information-Theoretic Analyses of Cone Mosaic Responses to Natural Scenes, ” Neural Comput., vol. 15(2), pp. 397-417, 2003.

J. H. van Hateren and D. L. Ruderman, “Independent Component Analysis of Natural Image Sequences Yields Spatio-Temporal Filters Similar to Simple Cells in Primary Visual Cortex,” Proc. Biological Sciences, vol. 265(1412), pp. 2315-2320, 1998.

H.B. Barlow, “Unsupervised learning,” Neural Computation, vol. 1, pp. 295–311, 1989.

S Deerwester, Dumais, S. T., Furnas, G. W., Landauer, T. K., & Harshman, R: *Indexing by latent semantic analysis*. Journal of the American Society for Information Science, 41(6), 391-407, (1990)

Isbell, Viola: “Restructuring sparse high dimensional data for effective retrieval” NIPS*11, 361-362 (1999)

T. Kolenda and L.K. Hansen: Independent Components in Text. In “Advances in Independent Component Analysis”, Perspectives in Neural Computing, Springer-Verlag p. 237-259 (2000).

J. Larsen, L. K. Hansen, T. Kolenda, F. Å. Nielsen: *Independent Component Analysis in Multimedia Modeling*, Proc. of ICA2003, Nara Japan, 687-696, (2003)

K.W. Jørgensen, L.L. Mølgaard, L.K. Hansen: Unsupervised Speaker Change Detection for Broadcast News Segmentation Eusipco, 2006 Florence, Italy, (2006).

L.L. Mølgaard, K.W. Jørgensen, L.K. Hansen, Castsearch - Context Based Spoken Document Retrieval. ICASSP, IEEE International Conference on Acoustics, Speech, and Signal Processing, Honolulu, Hawaii, (2007).