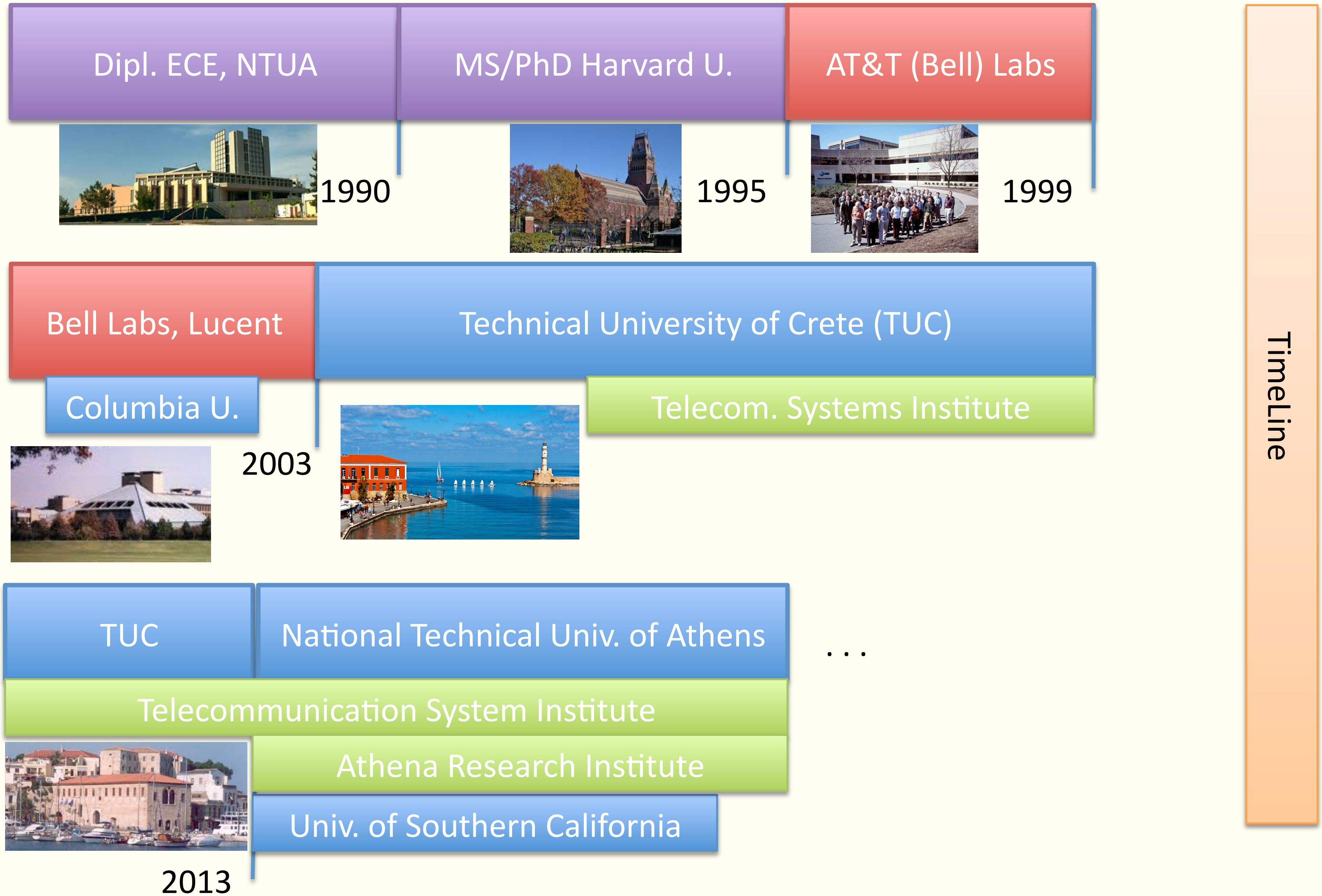
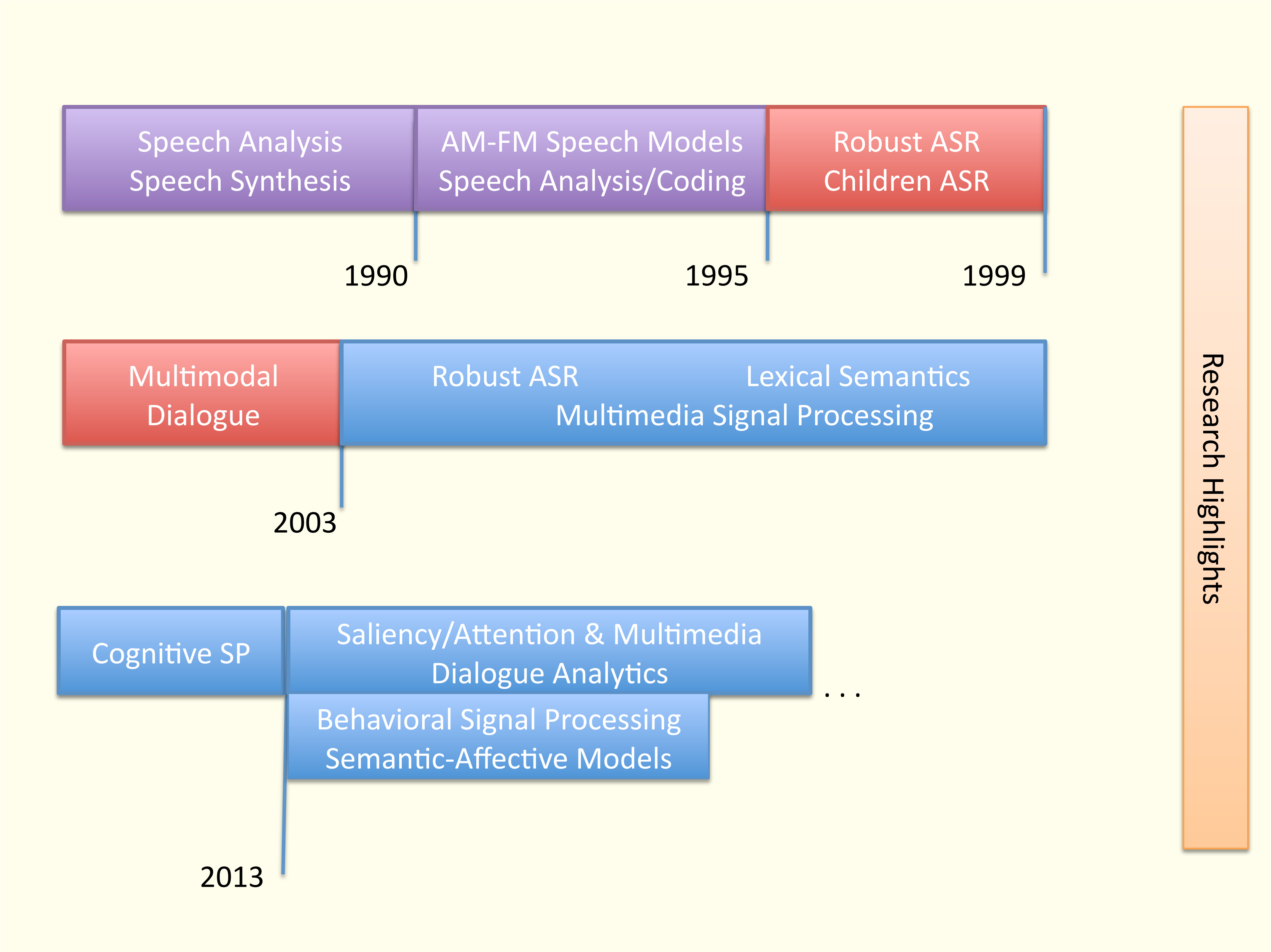
Semantic-Affective Models for Audio, Video and Text Processing

Alexandros Potamianos

National Tech. Univ. of Athens

Univ. of Southern California


Dipl. ECE, NTUA
MS/PhD Harvard U.
AT&T (Bell) Labs
1990
1995
1999
Bell Labs, Lucent
Technical University of Crete (TUC)
Columbia U.
Telecom. Systems Institute
2003
TUC
National Technical Univ. of Athens
. . .
Telecommunication System Institute
Athena Research Institute
Univ. of Southern California
2013
TimeLine


Speech Analysis
Speech Synthesis
AM-FM Speech Models
Speech Analysis/Coding
Robust ASR
Children ASR
1990
1995
1999
Multimodal
Dialogue
Robust ASR
Lexical Semantics
Multimedia Signal Processing
2003
Cognitive SP
Saliency/Attention & Multimedia
Dialogue Analytics
Behavioral Signal Processing
Semantic-Affective Models
. . .
2013
Research Highlights

Project Highlights

- DARPA Communicator Bell Labs 1999-2003
- HIWIRE EU-IST Robust ASR 2004-2007
- MUSCLE Network of Excellence on multimedia understanding 2005-2009
- Articulatory Speech Synthesis and Recognition GSRT 2008-2012
- PortDial EU-IST: resources for spoken dialogue systems 2012-2014
- CogniMuse GRST: multimedia semantics 2013-2016
- SpeDial EU-IST: spoken dialogue analytics 2013-2016
- BabyAffect GRST: language acquisition for autistic/TD children 2014-2016

Information Society Technologies

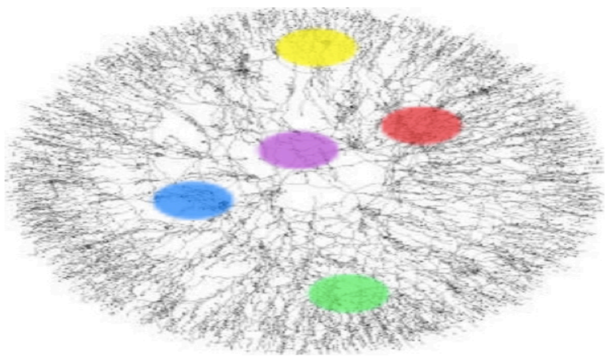
MINISTRY OF EDUCATION, LIFELONG LEARNING AND RELIGIOUS AFFAIRS
GENERAL SECRETARIAT FOR RESEARCH AND TECHNOLOGY

PortDial

www.muscle-noe.org

Research Highlights

- Affective analysis and classification of generic audio
- Emotion tracking of movies
- Salience/Attention models for movie summarization
- Cognitively-motivated semantic models/networks
- Low-dimensionality semantic representations



Outline

- Motivation
- Affective Modeling
  - Affective Classification of Audio Clips
  - Affective Tracking of Movies
- Multimedia and Cognition
  - Saliency and Attention
  - Application to movie summarization
- Semantic-Affective Models
  - Semantic similarity and DSMs
  - Affective text models

List of Open Questions

1. How are concepts, features/properties, categories, actions represented?
2. How are concepts, properties, categories, actions combined (compositionally)?
3. How are judgements (classification/recognition decisions) achieved?
4. How is learning and inference (especially induction) achieved?

Answers should fit evidence by psychology and neurocognition!

Three Solutions

- Symbolic
  - cognition is a Turing machine
  - computation is symbol manipulation
  - rule-based, deterministic (typically)
- Associationism, especially, connectionism (ANNs)
  - brain is a neural network
  - computation is activation/weight propagation
  - example-based, statistical, unstructured (typically)
- Conceptual
  - intermediate between symbolic and connectionist
  - concepts are represented as well-behaved (sub-)spaces
  - computation tools: similarity, operators, transformations
  - hierarchical, semi-structured

Properties of the Three Approaches

- Symbolic
  - Good for high-level cognitive computations (math)
  - Poor generalization power
  - Too expensive and slow for most cognitive purposes
- Conceptual
  - Excellent generalization power (intuition, physics)
  - Good for induction and learning; geometric properties (hierarchy, low dim., convex) guarantee quick convergence
  - Properties and actions defined as operators/translations
  - Still too slow for some survival-dependent decisions
- Connectionist (machine learning)
  - General-purpose, extremely fast and decently accurate
  - Computational sort-cuts create cognitive biases
  - Poor generalizability power due to high dimensionality and lack of crisp semantic representation

Properties of the Three Approaches

Property	Symbolic	Conceptual	Connectionist
cognitive speed	very slow	slow	fast
machine speed	very fast	pretty fast	fast
cognitive accuracy	good	good	decent
machine accuracy	decent	good	good
dimensionality	high	low	high
representation	flat	hierarchical	distributed
interpretability	excellent	good	low
determinism	high	medium	low
reasoning (all data)	good	good	decent
compositionality	good	good	decent
induction/learning	poor	excellent	average

Representation Learning

- Properties of a classifier with good generalization properties [Bengio et al 2013]:
  - Low-dimensionality/Sparseness
  - Distributed representations/hierarchy
    - Depth and abstraction
  - Shared factors across tasks
- Examples: auto-encoders, manifolds, deep neural nets …
- How to induce these properties in your classifiers:
  - Include as regularization term in training classifier criterion
  - Include properties directly in classifier design
  - Go deep and pray (dirty neural net tricks)

Our Goal

Cognitively-motivated semantic models

- Foreground-background classification using attention/ saliency
- Emphasis on induction not classification
- Associations not probabilities/distance
- Mappings between layers
- Hierarchical manifold models not metric spaces
- Multimodal not unimodal

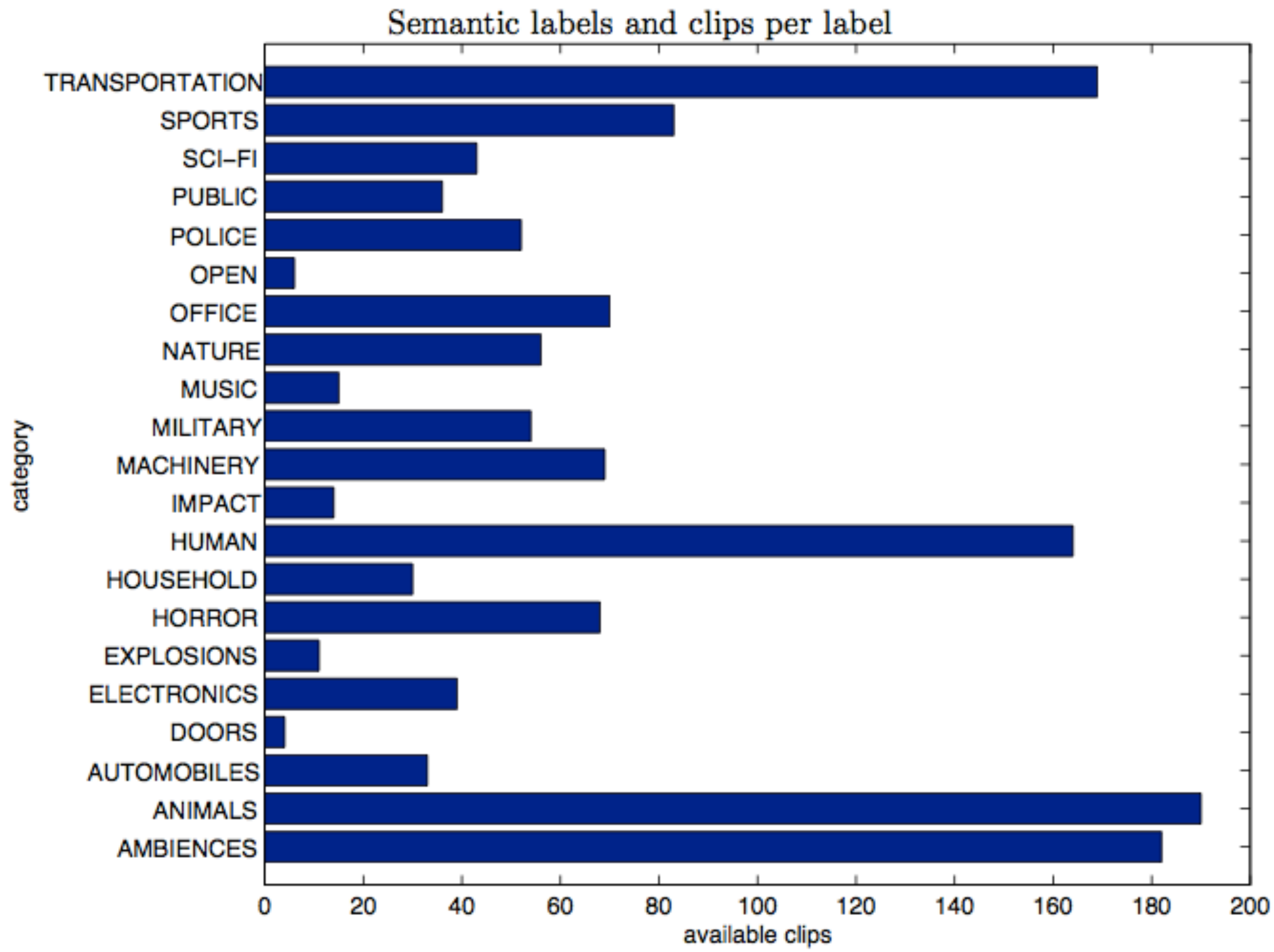
Part I: Affective Modeling of Multimedia

Affective Classification of Generic Audio Clips using Regression Models

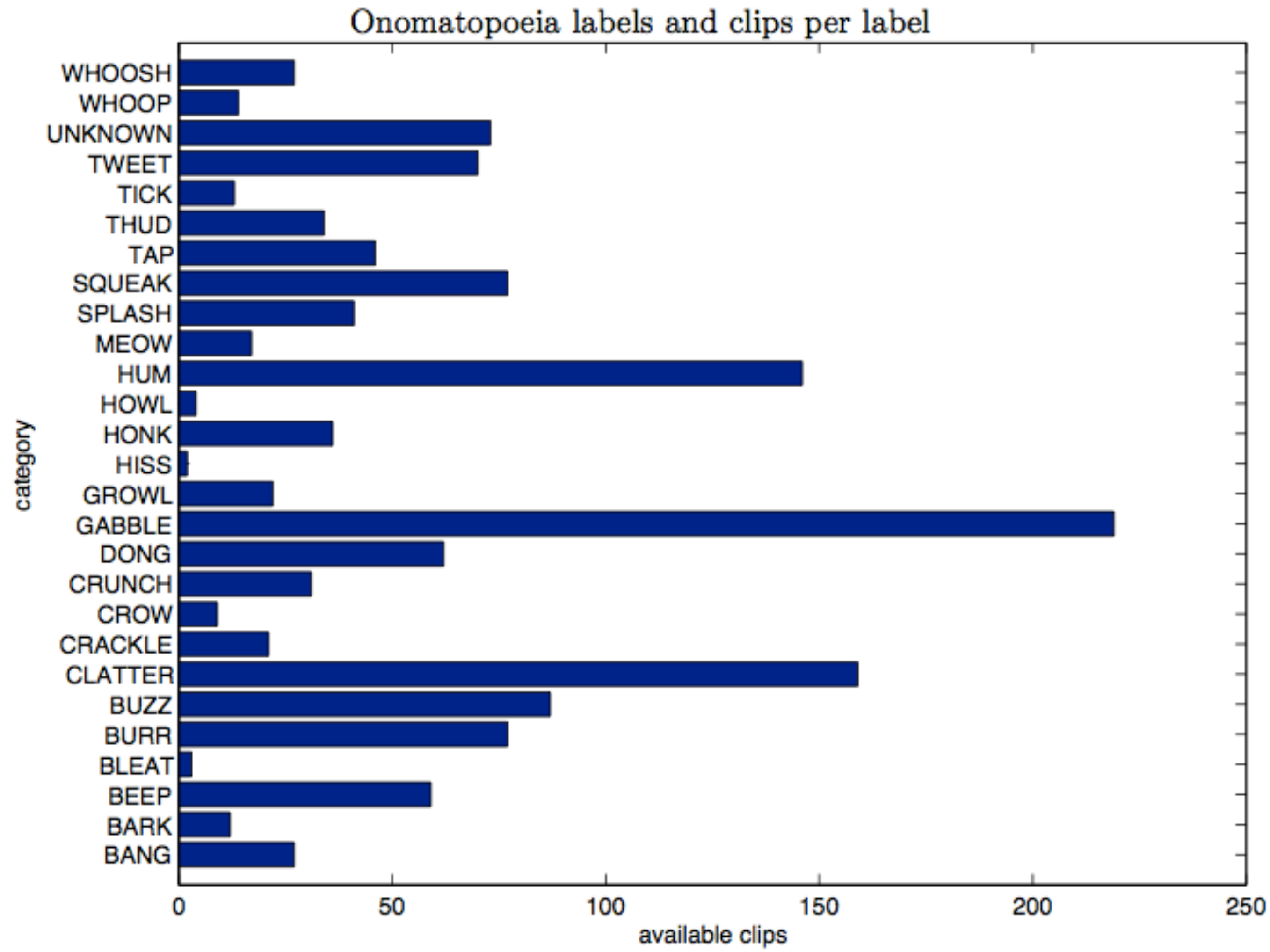
N. Malandrakis, S. Sundaram, A. Potamianos

InterSpeech 2013

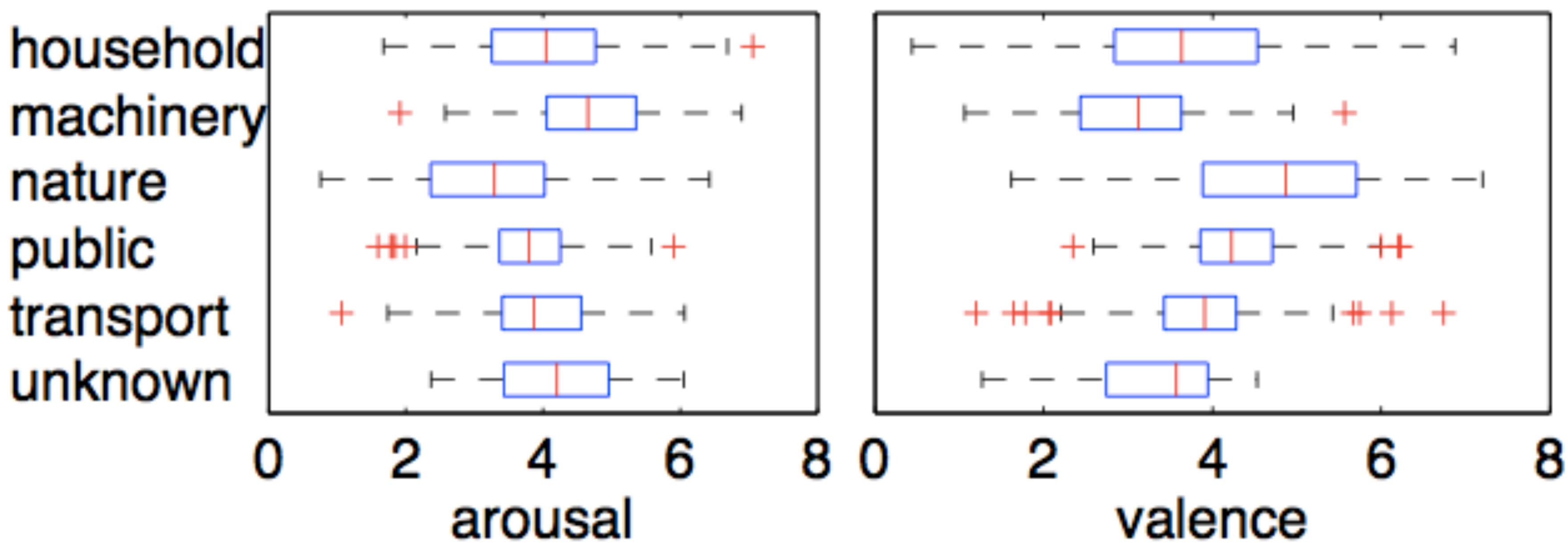
Semantics of Generic Audio I



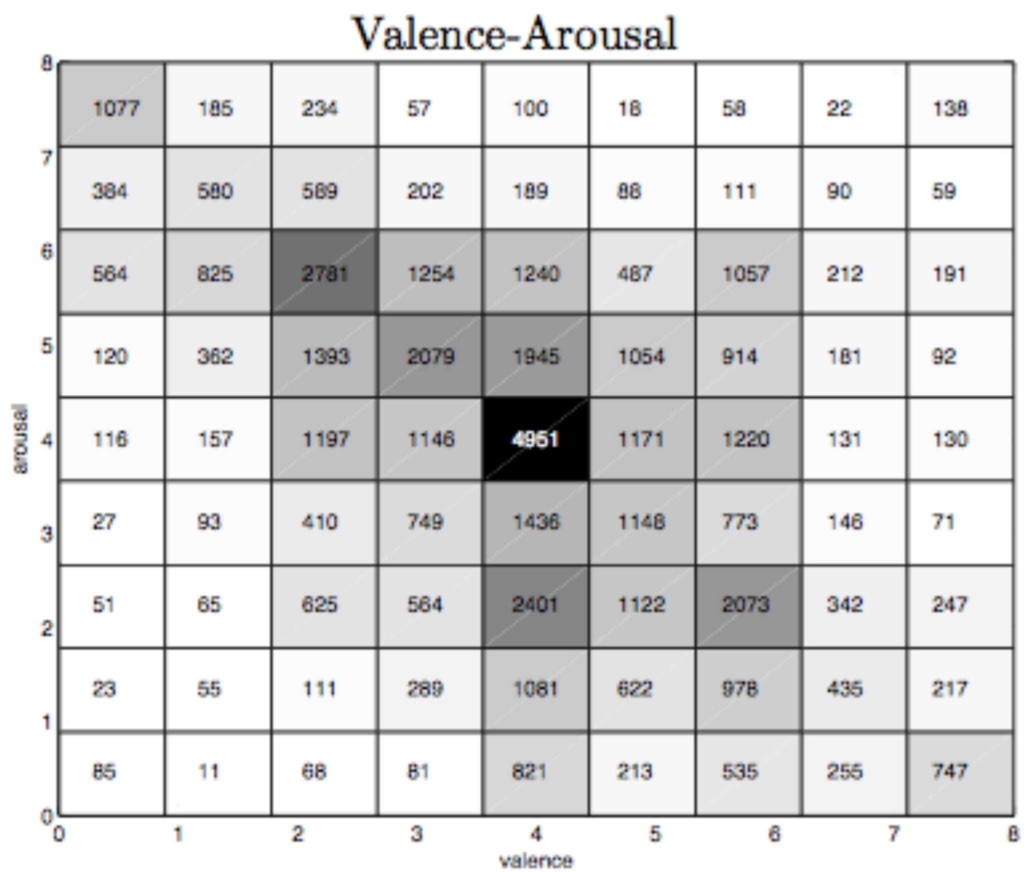
Semantics of Generic Audio II



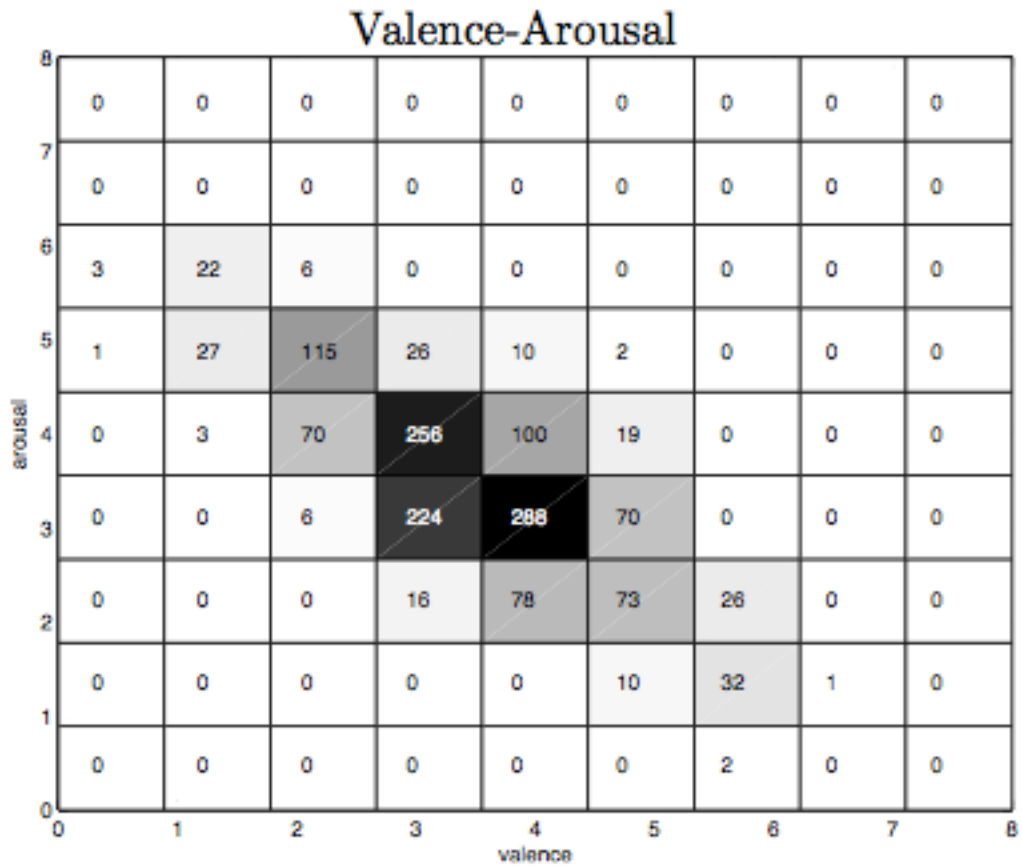
Overall affective characterization



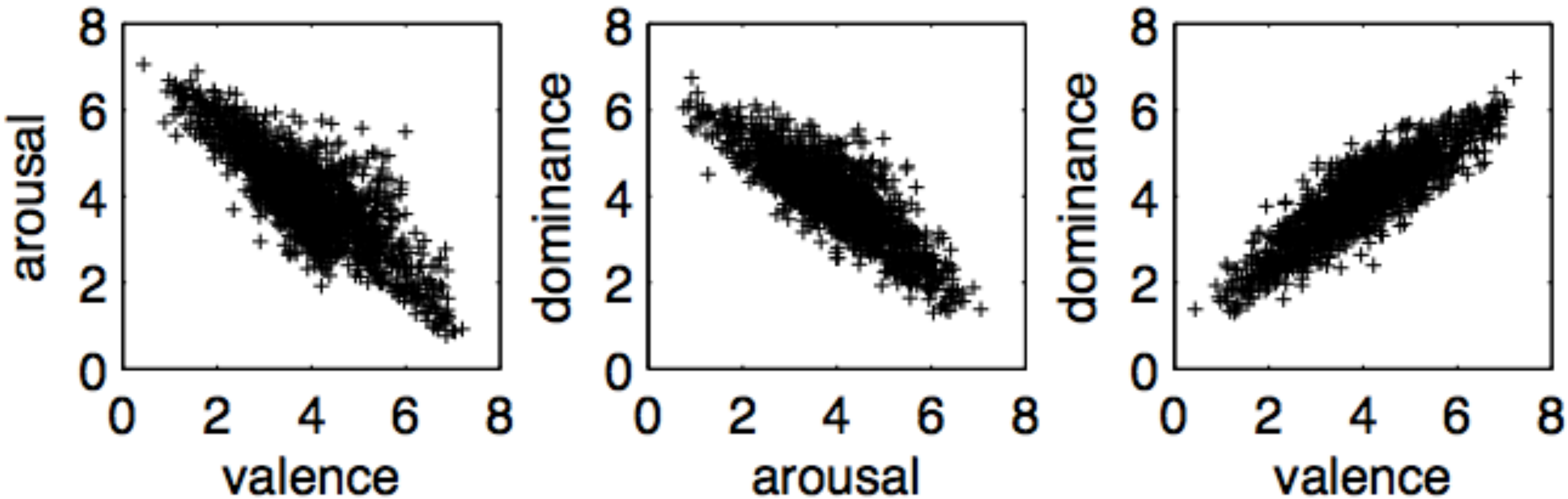
Distribution of All Ratings



Distribution of Clip Average Ratings



3D Affective space correlations



Inter-annotator agreement

Inter-annotator agreement			
Metric	Arous.	Valen.	Domn.
avg. pairwise correlation	0.52	0.55	0.16
avg. pairwise mean abs. dist.	2.02	1.84	2.32
Krippendorff's alpha (ordinal)	0.39	0.47	0.11
Krippendorff's alpha (interval)	0.39	0.46	0.10
Agreement with the ground truth			
Metric	Arous.	Valen.	Domn.
avg. correlation	0.55	0.60	0.41
avg. mean abs. dist.	1.42	1.18	1.36

Frame level vs Long-Term Features

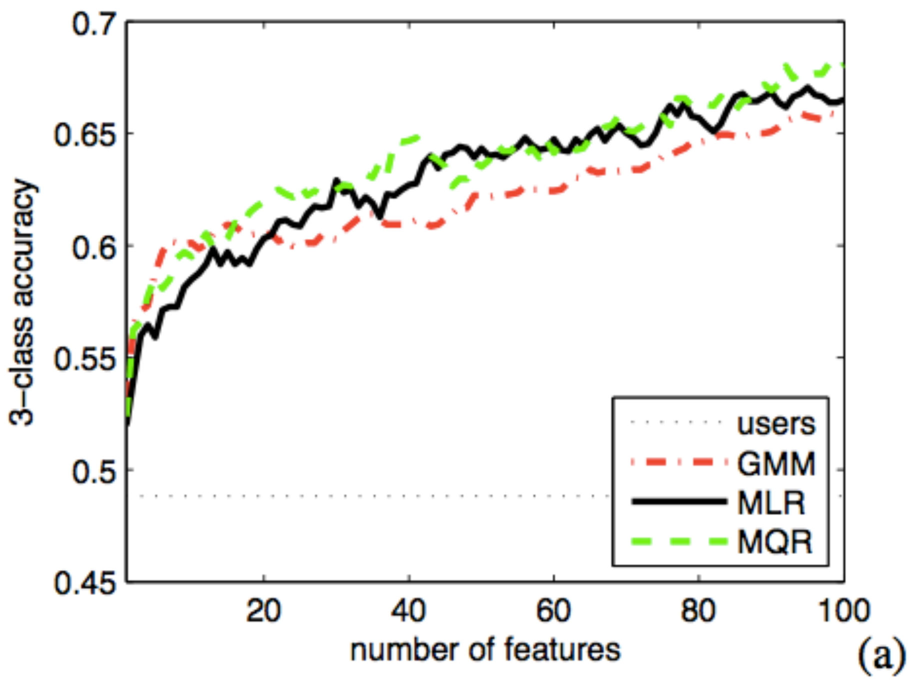
Scope	Low Level. Descr.	Arous.	Valen.	Domn.
frame level	chroma + Δ	0.41	0.45	**0.43**
	log Mel power + Δ	0.44	0.48	0.44
	MFCC + Δ	0.45	0.44	0.43
long term	chroma + Δ	0.41	**0.46**	0.42
	log Mel power + Δ	**0.46**	**0.49**	**0.46**
	MFCC + Δ	**0.48**	**0.48**	**0.45**

Feature Selection

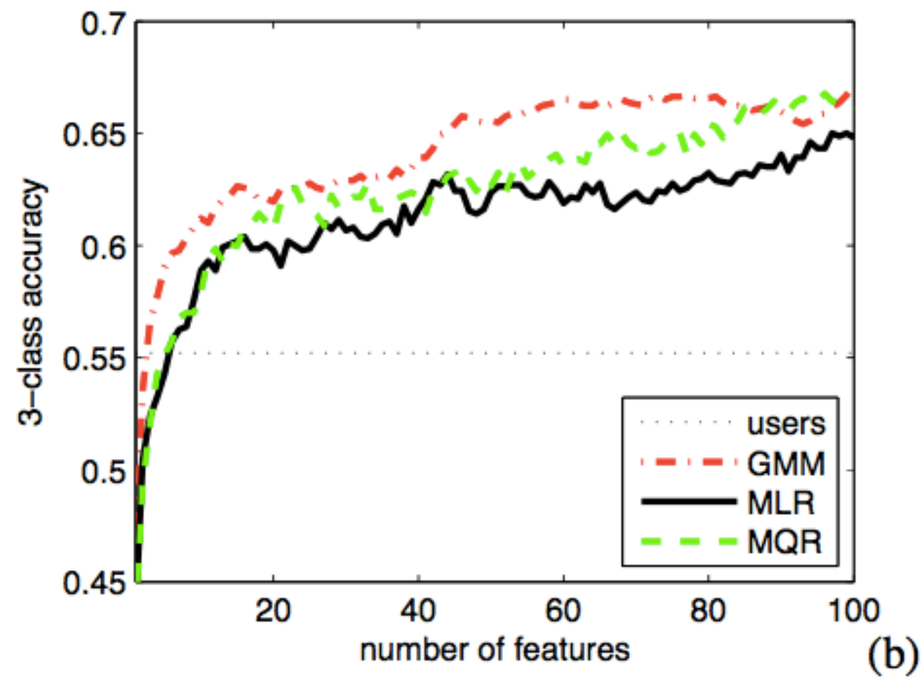
Model	# of features	Arous.	Valen.	Domn.
Users	-	0.55	0.60	0.41
MLR Regression Model	10	0.70	0.67	0.63
	20	0.72	0.70	0.65
	30	0.74	0.71	0.67
	40	0.75	0.72	0.68
	50	**0.75**	**0.73**	**0.69**

3-class Classification Accuracy

Arousal



Valence



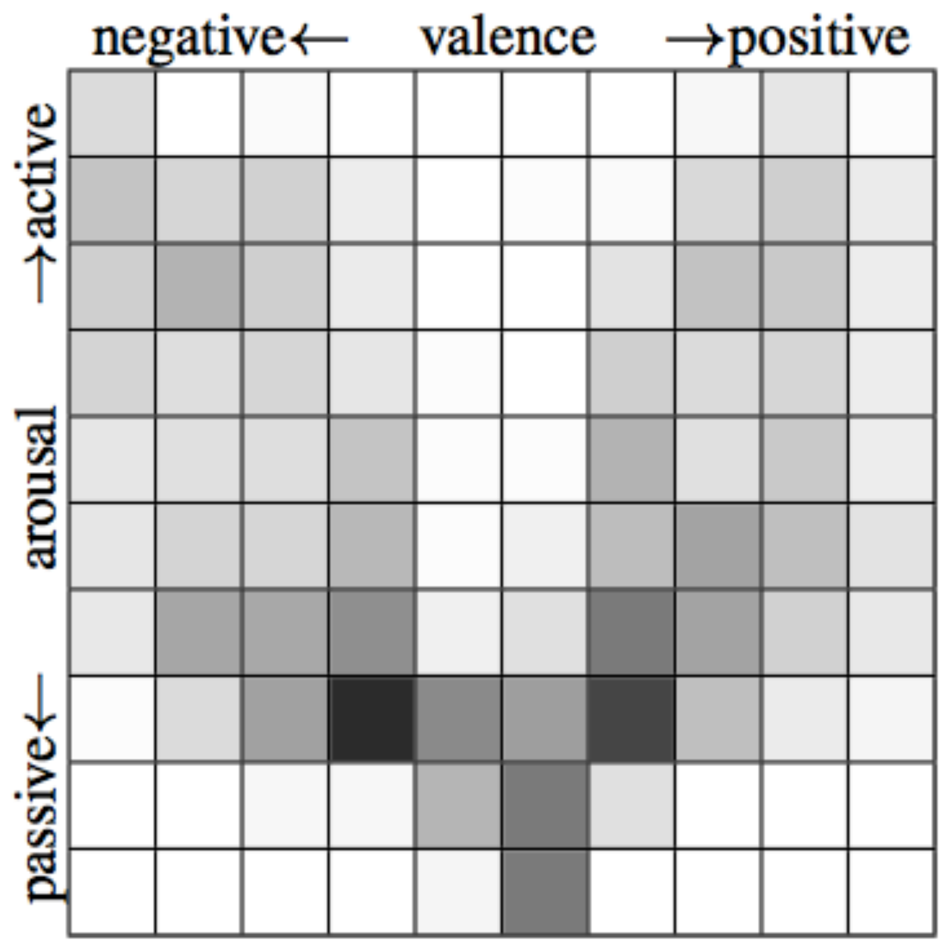
A Supervised Approach to Movie Emotion Tracking

N. Malandrakis, A. Potamianos, G. Evangelopoulos, A. Zlatintsi

ICASSP 2011

Example Frames

Arousal vs Valence Labeled Data



Features and Models

- Continuous-time modeling using HMM models
- Language model used for smoothing
- Features used:

Valence	audio	12 MFCCs and C0, plus derivatives
	video	maximum color value
	video	maximum color intensity
Arousal	audio	12 MFCCs and C0, plus derivatives

Results: Frame Confusion Matrix

Arousal

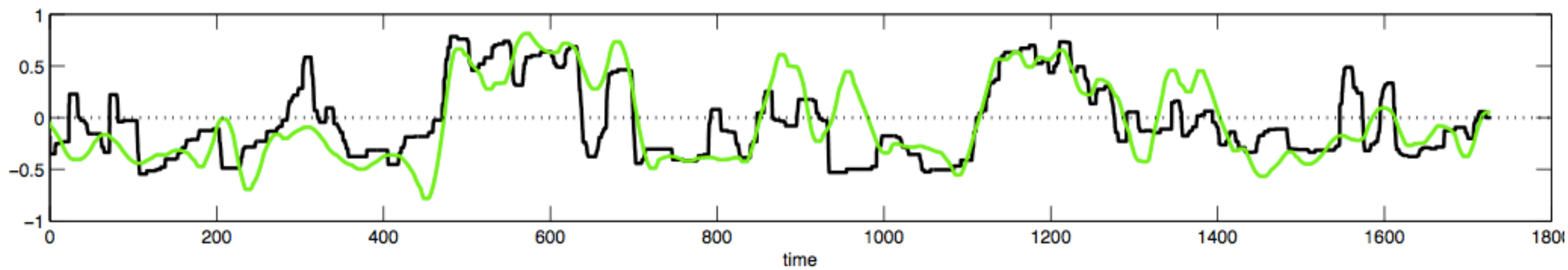
actual \ predicted	passive←			predicted			→active
→active	3	4	10	6	9	17	51
	5	9	14	13	13	21	25
	6	13	23	16	9	21	12
actual	11	13	27	22	10	10	7
	11	18	29	19	11	9	3
passive←	17	16	28	18	8	10	3
	24	18	23	14	6	13	2

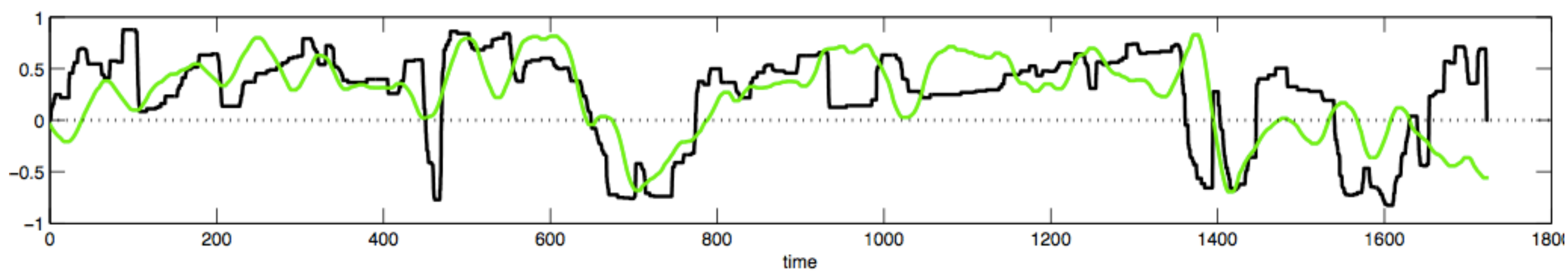
Valence

actual \ predicted	negative←			predicted			→positive
→positive	2	6	7	10	25	34	16
	5	5	10	13	20	29	18
	3	6	15	18	20	23	15
actual	6	17	26	24	16	8	3
	8	26	30	20	8	6	2
negative←	13	25	25	15	9	6	7
	18	30	22	11	6	9	4

Continuous-Time Emotion Tracking

Arousal



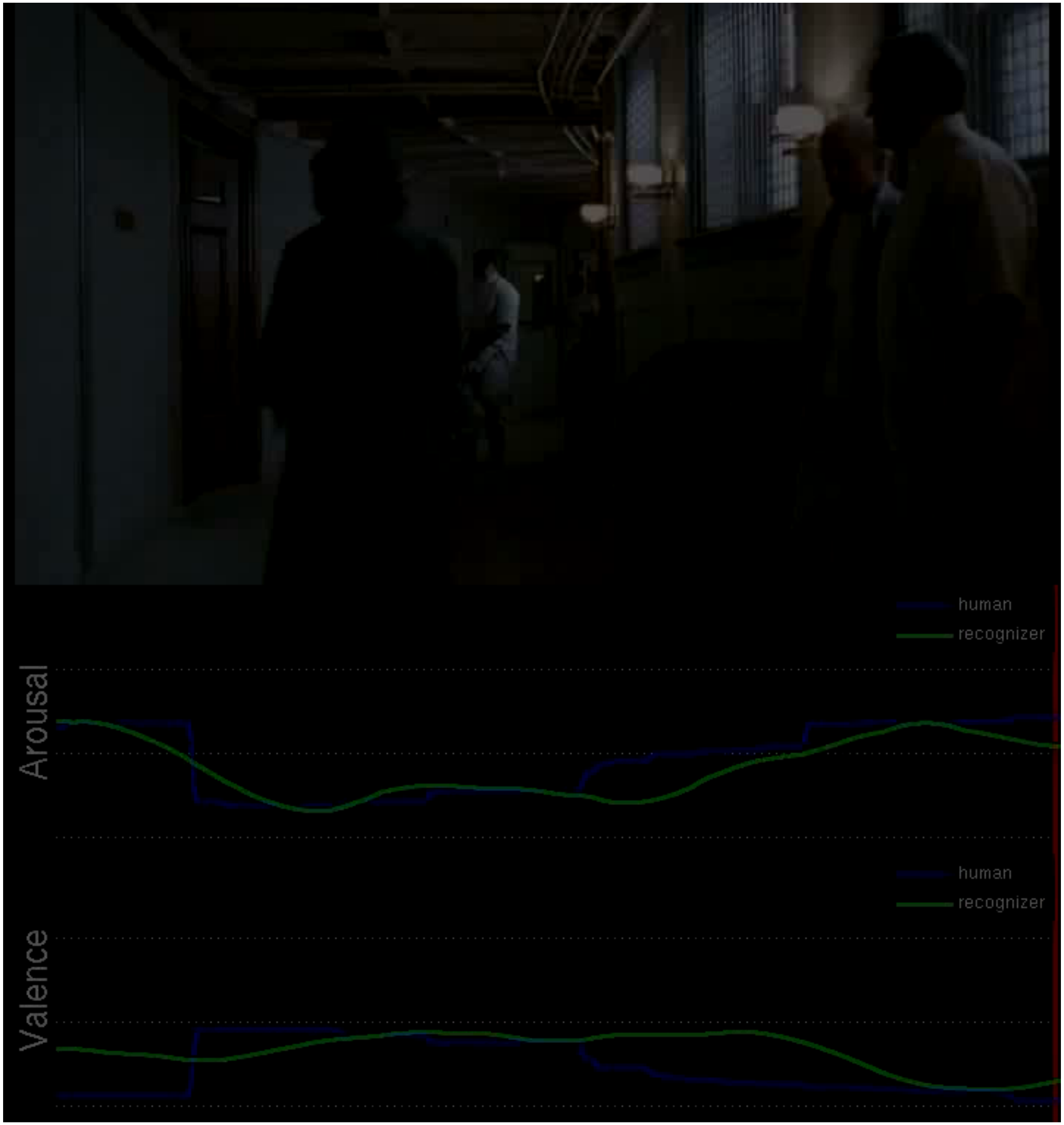


Valence

Affective tracks:
Arousal & Valence

Green– Machine

Blue – Human
Annotators (average)



Discussion

- Affective analysis of generic audio using frame-level features and their statistics
- Affect of movies fusing multimodal cues
- Hard to draw general conclusions about feature selection
  - No universal features (except MFCCs!?)
- A detection-based approach for audio processing?

Saliency, Attention and Summarization in Movies

Cognition and Attention

- What grabs our attention?
  - Salient events
- Attention and Perception:
  - A **simple** perceptual algorithm
  - Quickly identify relevant (to survival) information
  - Bottom-up selectional attention: features extracted via low level signal processing
  - Fusion of top-down and bottom-up attention
- The attention/saliency relationship is used in multimedia production

What

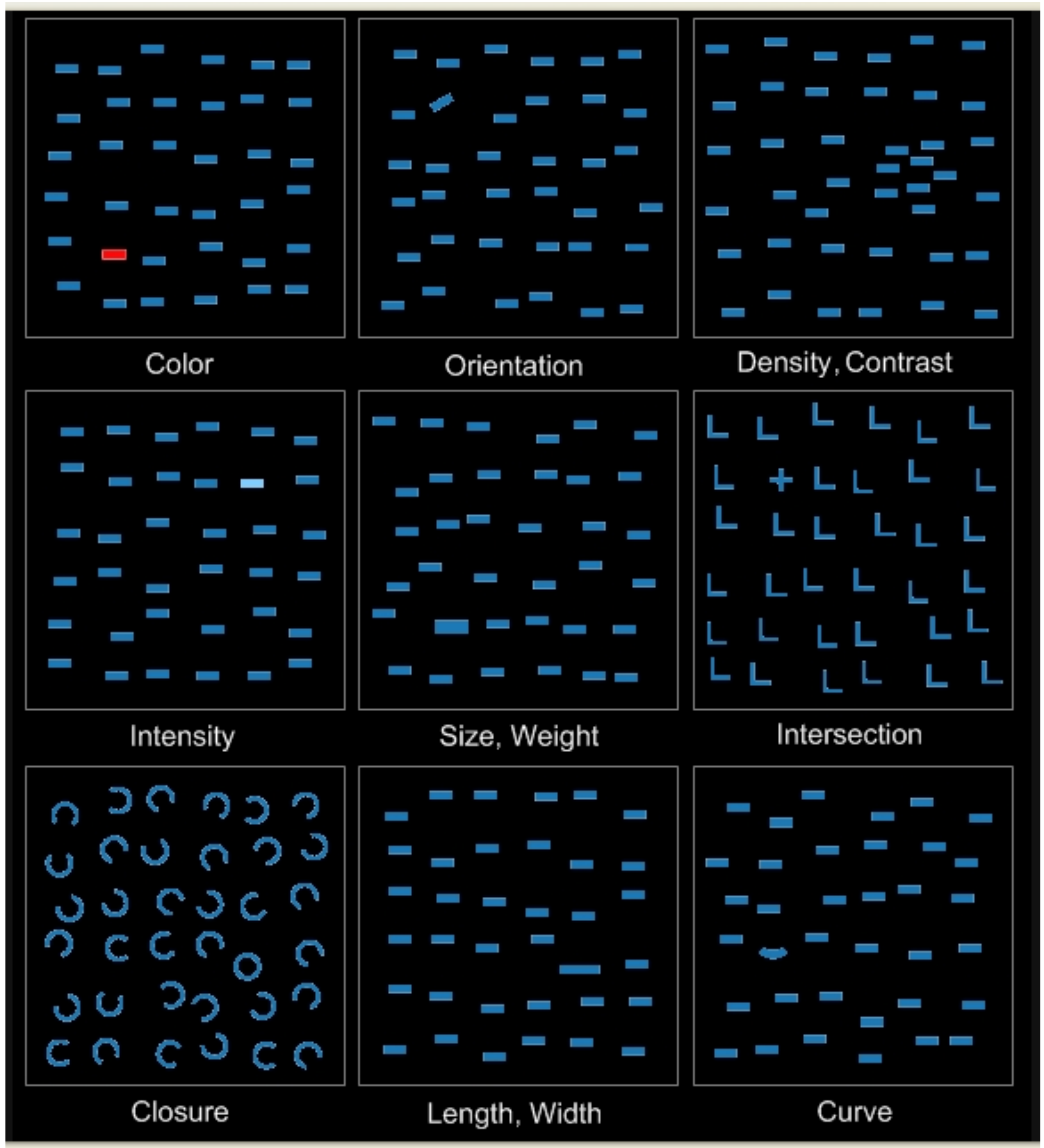
Grabs

Your

Attention

in an

Image?

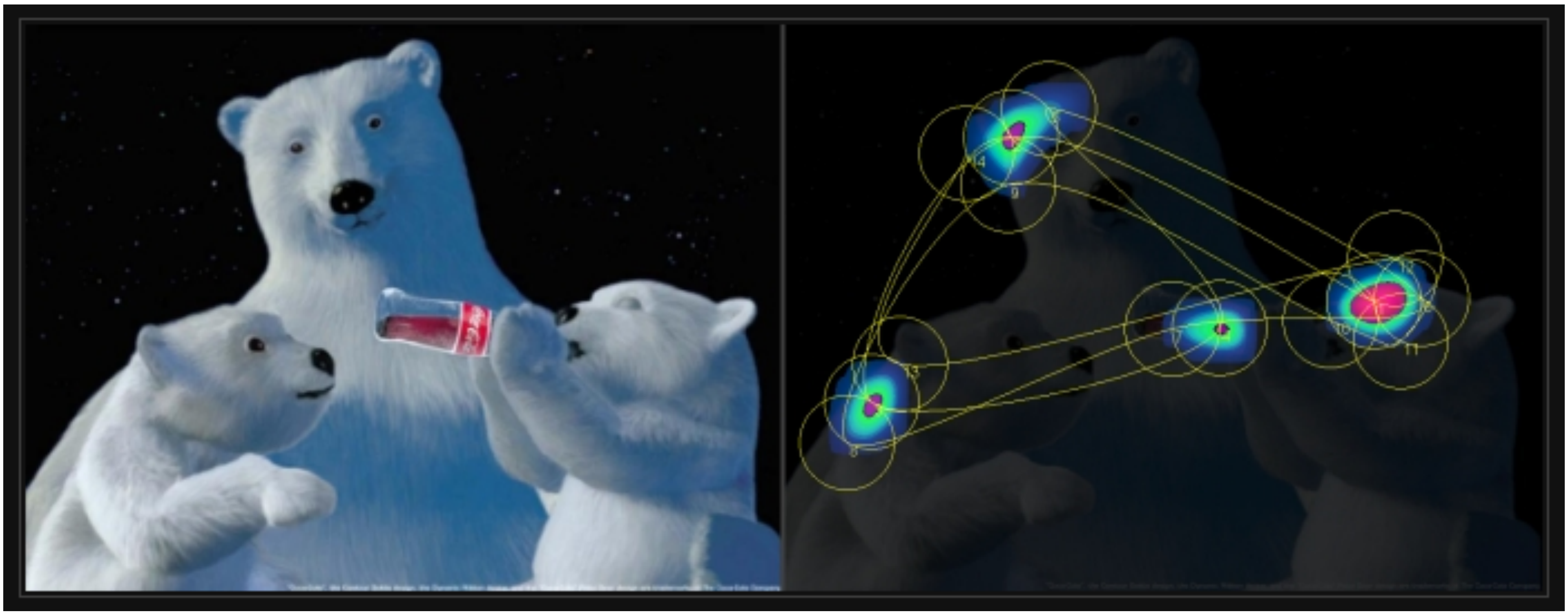


from http://www.feng-gui.com

Attention and Saliency

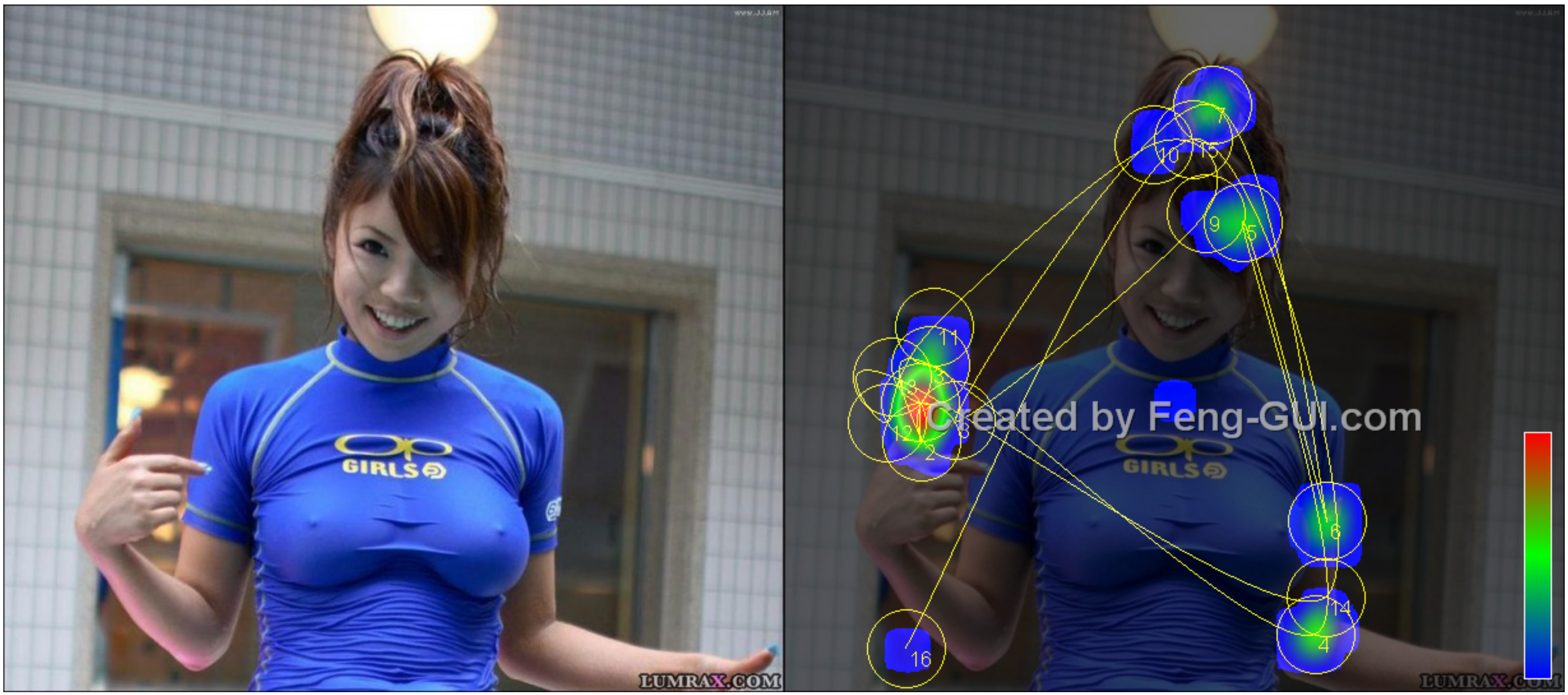
- Audio: rhythm, energy, change of frequency content
- Images over time (video): motion (direction, velocity), flicker
- Such low level features capture about 60-80% of "events" in each modality
- How do we capture the rest?
  - Multimodality (up to 90%)
  - Semantics (top-down selectional attention)

Attention Models: Good Example



example from http://www.feng-gui.com

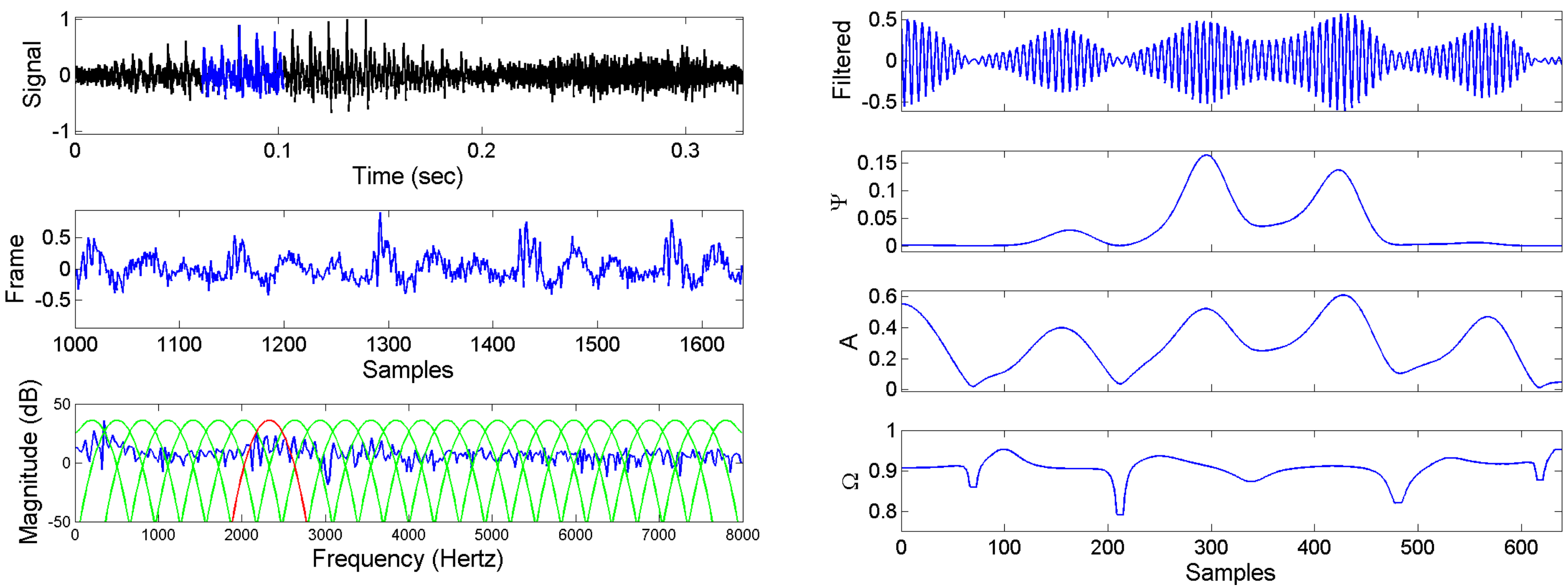
Attention Models: Bad Example



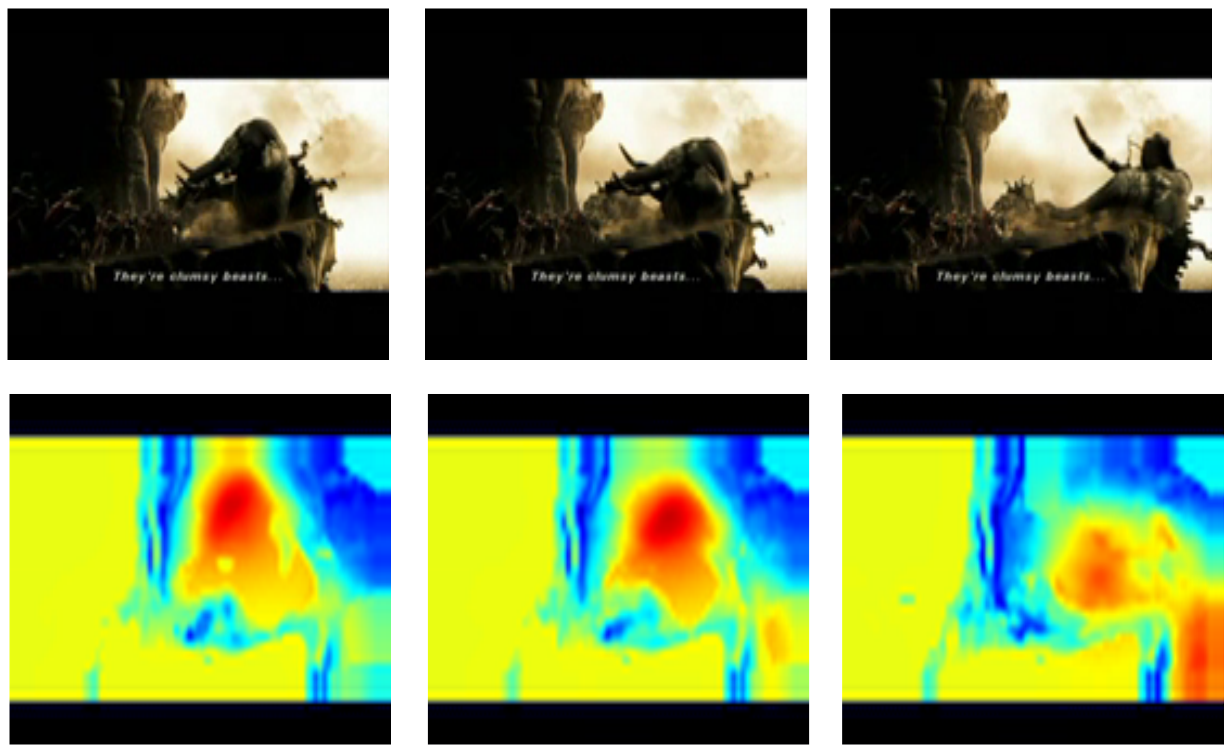
Attention Models and Saliency

* Attention model of video streams
* Saliency measures:
  - Aural: energy of multi-frequency band features
  - Visual: multi-scale intensity, color and motion
  - Text: part of speech assignments
* Fusion on a single audio-visual-text saliency metric

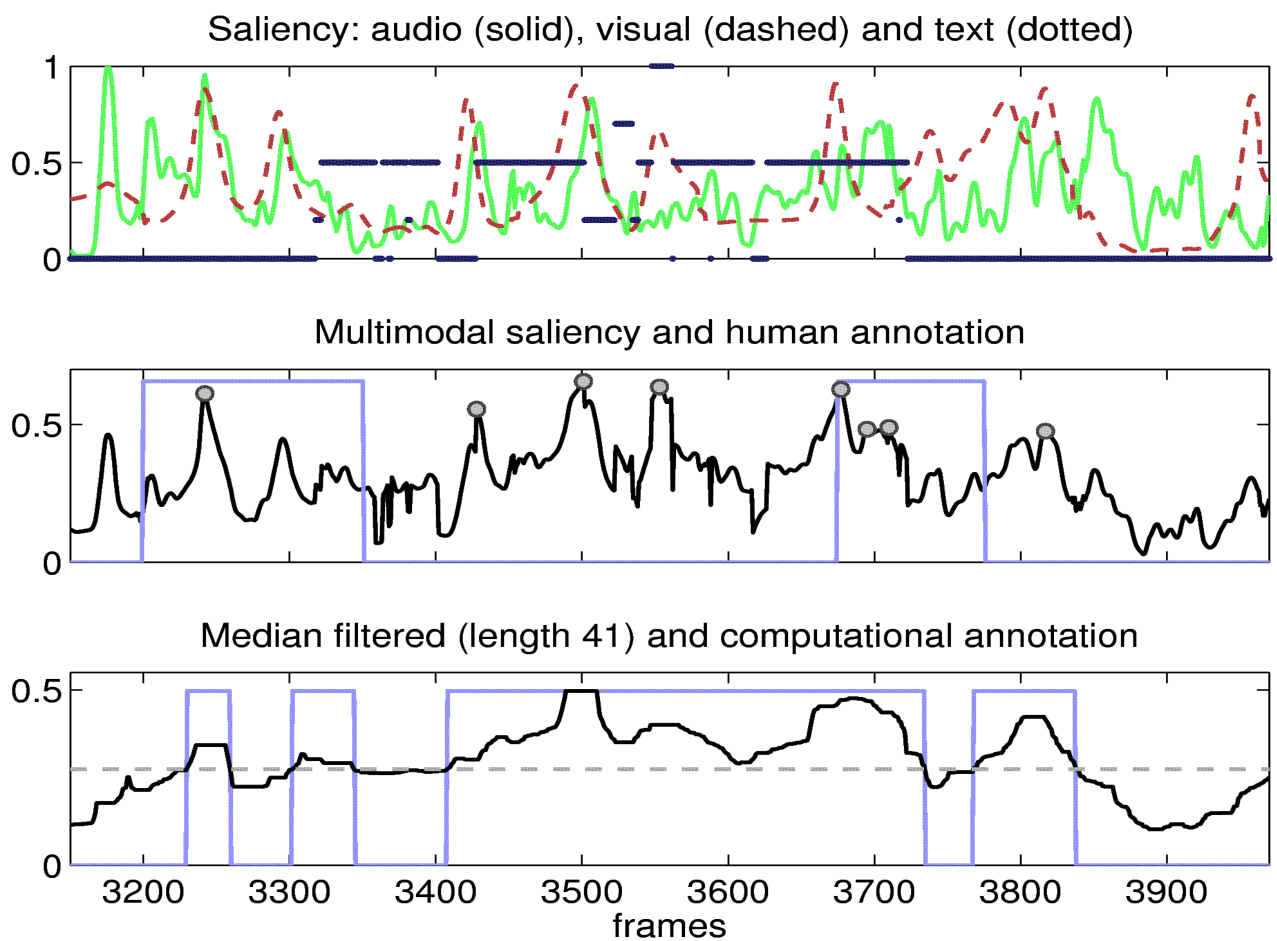
Audio Saliency Features



Visual Saliency



AVT Salience via Linear Fusion



Example: x2 compression

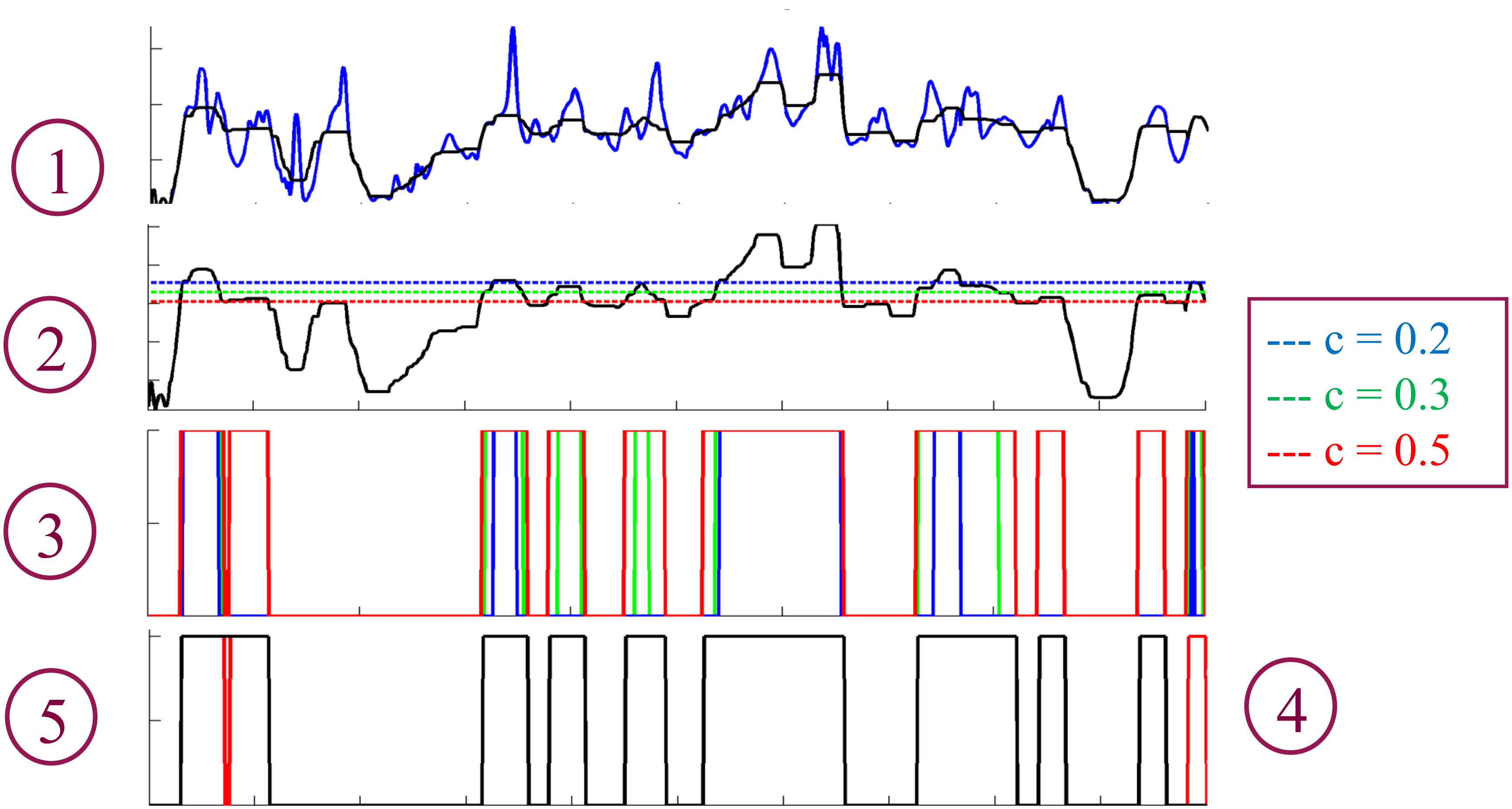
AV Key Frames: 300

Movie Summarization Algorithm

1. Filter: AVSC with median of length $2M + 1$.
2. Threshold choice
3. Selection: segments
4. Reject: segments shorter than N frames
5. Join: segments less than K frames apart
6. Render: Linear overlap-add on L video frames and audio

Evaluation: $M = N = 20$, $K = L = 10$ (videos at 25 fps).

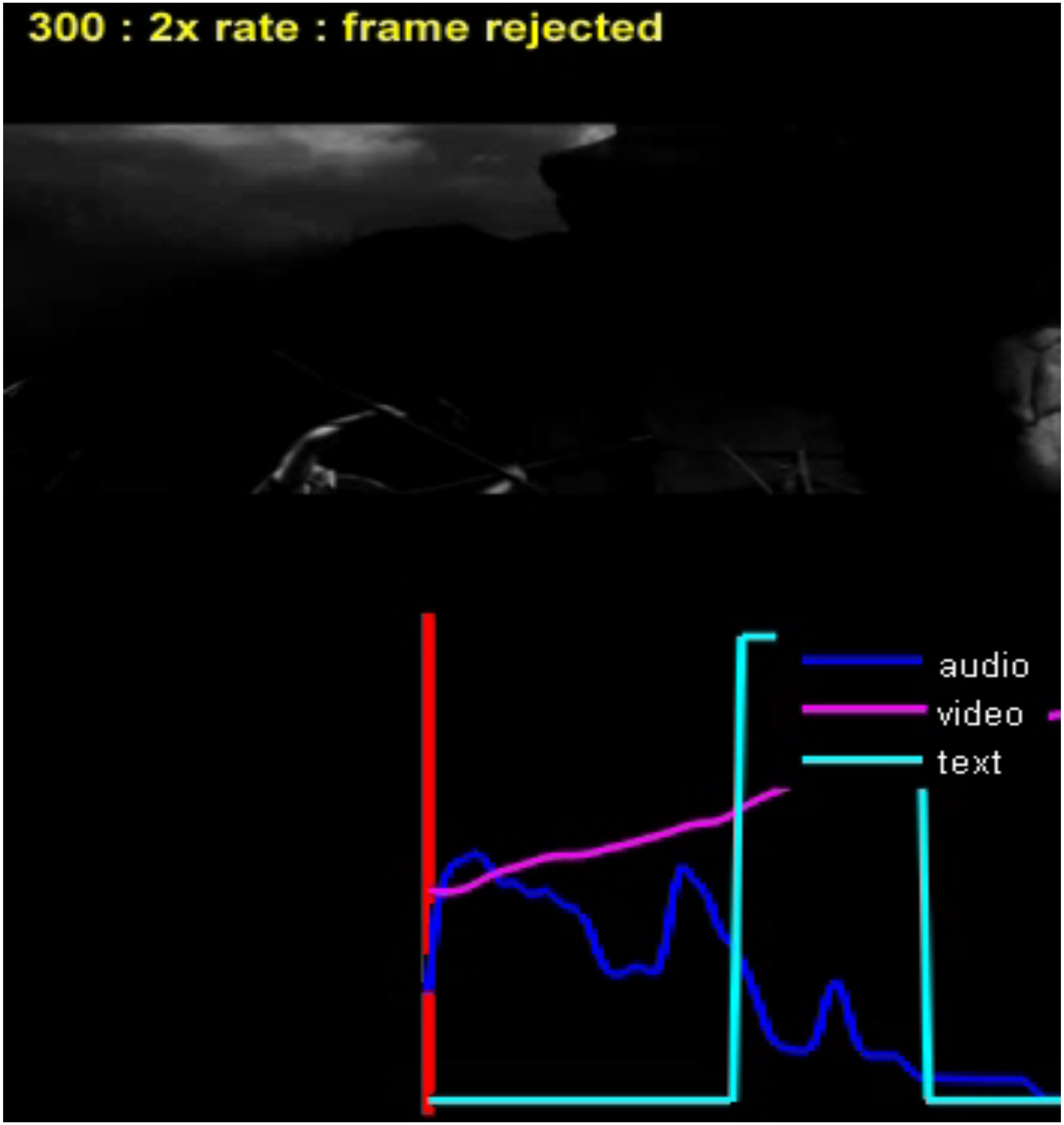
Movie Summarization Algorithm (2)



Summary annotated with AVT Saliency

Grey – Rejected

Color- Accepted in summary



Discussion

- Low-level selectional attention can be modeled using
  - Low level feature detectors
  - Fusion of detectors across modalities
  - Can capture up to 95% of semantics
- Ongoing work
  - Attentional mechanisms in audio beyond energy
  -]Text saliency
  - Semantics – Plot Analysis

Part III: Semantic Representations

Acknowledgements

- Elias Iosif, Kelly Zervanou,Maria Giannoudaki: Semantic similarity computation, semantic networks
- Nikos Malandrakis: Affective models for text and multimedia
- Georgia Athanasopoulou: Metric semantic spaces
- Shri Narayanan (USC): Affective modeling of dialogue interaction

References

[1] E. Iosif and A. Potamianos. 2010. "Unsupervised semantic similarity computation between terms using web documents". IEEE Transactions on Knowledge and Data Engineering.
[2] N. Malandrakis, A. Potamianos, E. Iosif, S. Narayanan. 2011. "Kernel methods for affective lexicon creation". Proc. Interspeech.
[3] — . 2011. "EmotiWord: Affective Lexicon Creation with Application to Interaction and Multimedia Data". Proc. of MUSCLE workshop.
[4] E. Iosif and A. Potamianos. 2012. "Semsim: Resources for normalized semantic similarity computation using lexical networks". In Proc. LREC.
[5] N. Malandrakis, E. Iosif, A. Potamianos. 2012. "DeepPurple: Estimating Sentence Semantic Similarity using N-gram Regression Models and Web Snippets". In Proc SemEval (collocated with NAACL-HLT).
[6] E. Iosif and A. Potamianos. 2013. "Similarity computation using semantic networks created from web-harvested data". Natural Language Engineering.
[7] N. Malandrakis, A. Potamianos, E. Iosif and S. Narayanan. 2013. "Distributional Semantic Models for Affective Text Analysis". IEEE Transactions on Audio, Speech and Language Processing.

Problem Definition

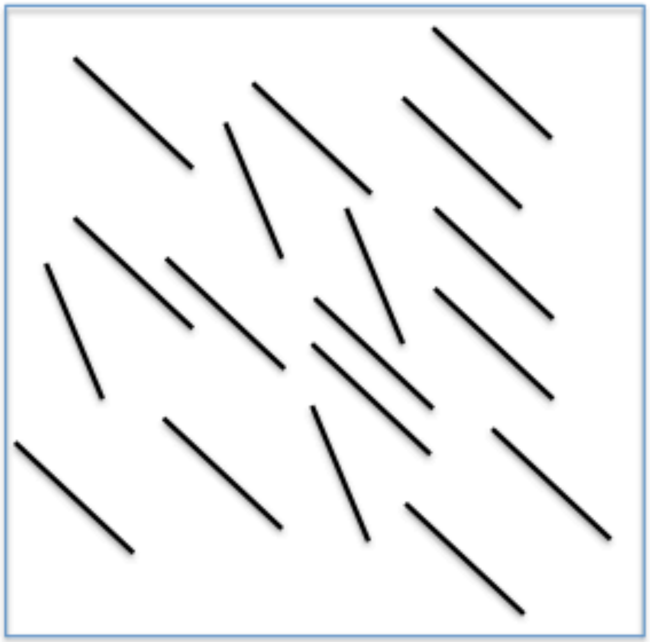
- Semantic Similarity Computation
  - Given a pair of words or terms (w_i, w_j)
  - Compute semantic similarity between them $S(i, j)$
- Related tasks
  - Phrase or sentence level semantic similarity
  - Strength of associative relation between words
  - Affective score (valence) of words and sentences
- Motivation
  - Organizing principle of human cognition
  - Building block of machine learning in NLP/semantic web
  - Entry point for the semantics of language

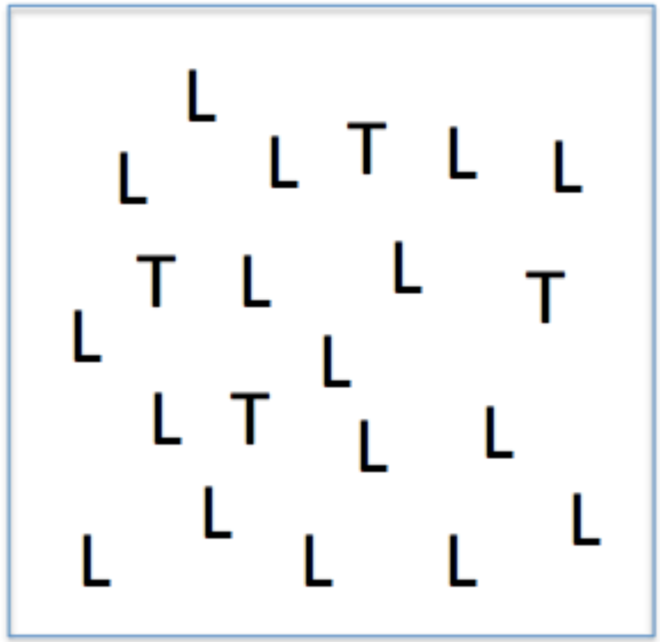
System 1 vs System 2

- Using Kahneman's (and others) formalism:
  - System 1 (intuition): generates
    - impressions, feelings, and inclinations
  - System 2 (reason): turns System 1 input into
    - beliefs, attitudes, and intentions
- Associative relations reside in System 1
- But where do semantic relations reside?

Example

- Example from vision: system 1 vs system 2



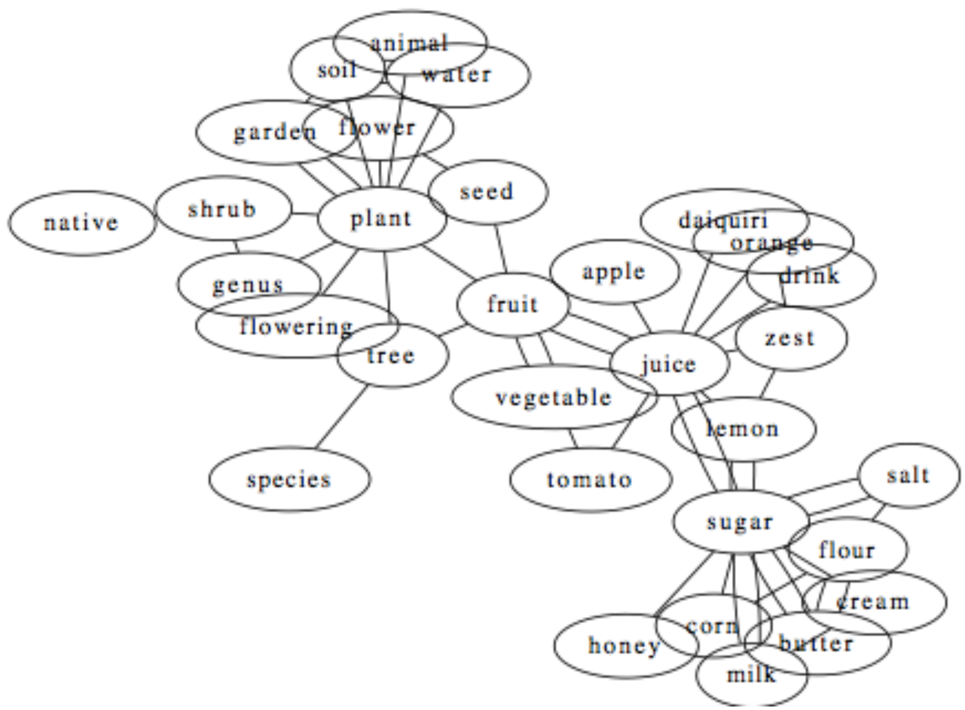


Main approaches of lexical semantics

- Word are associated with feature vectors
  - crisp, parsimonious representation of semantics
- Distributional semantic models (DSMs)
  - Semantic information extracted from word frequencies
  - Estimate co-occurence counts of word pairs or triplets
  - Estimate statistics of word context vectors
- Semantic networks
  - discovery of new relations via systematic co-variation
  - robust estimates – smoothing corpus statistics over network
  - rapid language acquisition

Example of Semantic Network

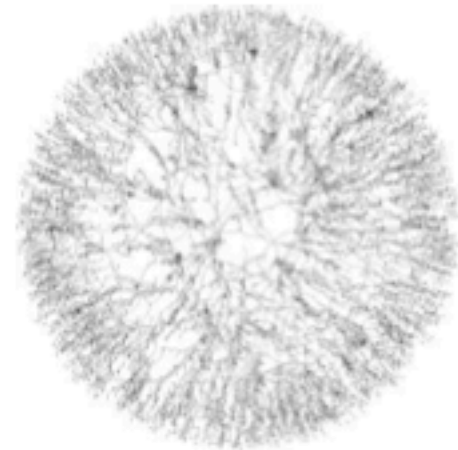
- Linked nodes: lexicalized senses and attributes
  - Informative for semantic similarity computation
- Computation of structural properties, e.g., cliques



Proposed semantic similarity two-tier system

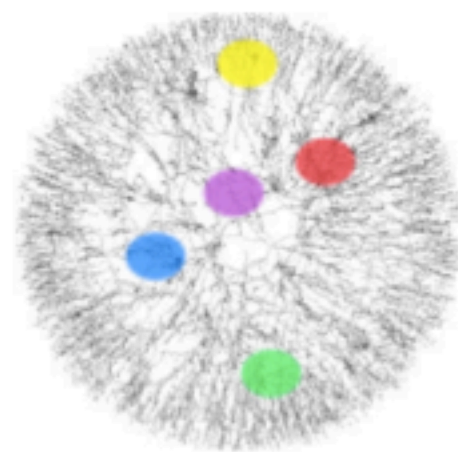
- Unifies the three approaches
- Fuzzy vs explicit semantic relations
- Word senses vs words vs concepts
- A two tier system
  - An associative network backbone
  - Semantic relations defined as operations on network neighborhoods (cliques)
- Consistent with system 1 vs system 2 view
- Furthermore we believe that the
  - underlying network consists of word senses, and
  - is a low dimensional semi-metric space

Lexical Network - Semantic Neighborhoods



Lexical Network

- Undirected graph $G = (N, E)$
  - Vertices N: words in lexicon L
  - Edges E: word similarities



Semantic Neighborhoods

- For word i create subgraph G_i
- Select neighbors of i
  - Compute $S(i, j), \forall j \in L, i \neq j$
  - Sort j according to $S(i, j)$
  - Select $|N_i|$ top-ranked j

Semantic Neighborhoods: Examples

Word	Neighbors
automobile	**auto**, truck, **vehicle**, **car**, **engine**, bus, . . .
car	truck, **vehicle**, travel, service, **price**, **industry**, . . .
slave	slavery, beggar, **nationalism**, society, **democracy**, **aristocracy**, . . .
journey	**trip**, holiday, **culture**, **travel**, **discovery**, **quest**, . . .

- Synonymy
- Taxonomic: IsA, Meronymy
- Associative
- Broader semantics/pragmatics

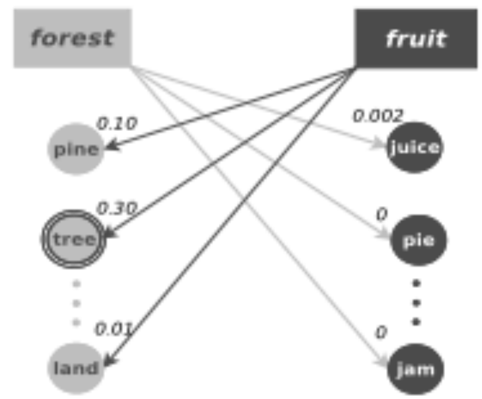
Semantic Sim. Computation: Sense Similarity

- Maximum sense similarity assumption *[Resnik, '95]*:
  - Similarity of words equal to similarity of their closest senses
  - If words are considered as sets of word senses, this is the "common sense" set distance
- Given words w_1, w_2 with senses s_{1i}, s_{2j}

$$S(w_1, w_2) = \max_{ij} S(s_{1i}, s_{2j})$$

Neighborhood-based Similarity Metrics: M_n

M_n metric: maximum similarity of neighborhoods



- Motivated by maximum sense similarity assumption
  - Neighbors are semantic features denoting senses
  - Similarity of two closest senses
- Select max. similarity: $M_n(\text{“}forest\text{”}, \text{“}fruit\text{”}) = 0.30$

Performance of net-based similarity metrics

- Task: similarity judgment on noun pairs
- Dataset: MC *[Miller and Charles, 1998]*
- Evaluation metric: Pearson's correlation wrt to human ratings

Dataset	Neighbor selection	Similarity computation	Metrics		
			$M_{n=100}$	$R_{n=100}$	$E^{\theta=2}_{n=100}$
MC	co-occur.	co-occur.	0.90	0.72	**0.90**
MC	co-occur.	context	**0.91**	0.28	0.46
MC	context	co-occur.	0.52	**0.78**	0.56
MC	context	context	0.51	0.77	0.29

Performance of web-based similarity metrics

- For MC dataset

Feature	Description	Correlation
context	AND queries	0.88
context	IND queries	0.55
context	IND queries: network	0.90

- Comparable to structured DSMs, WordNet-based approaches

Contributions

Proposed a language agnostic, unsupervised and scalable algorithm for semantic similarity computation

- No linguistic knowledge required, works from text corpus or using a web query engine
- Shown to perform at least as well as resource-based semantic similarity computation algorithms, e.g., WordNet-based methods

Motivation

- Affective text labeling at the core of many multimedia applications, e.g.,
  - Sentiment analysis
  - Spoken dialogue systems
  - Emotion tracking of multimedia content
- Affective lexicon is the main resource used to bootstrap affective text labeling
  - Lexica are currently of limited scope and quality

Goals and Contributions

Our goal: assigning continuous high-quality polarity ratings to any lexical unit

- We present a method of expanding an affective lexicon, using web-based semantic similarity
- Assumption: semantic similarity implies affective similarity.
- The expanded lexica are accurate and broad in scope, e.g., they can contain proper nouns, multi-word terms

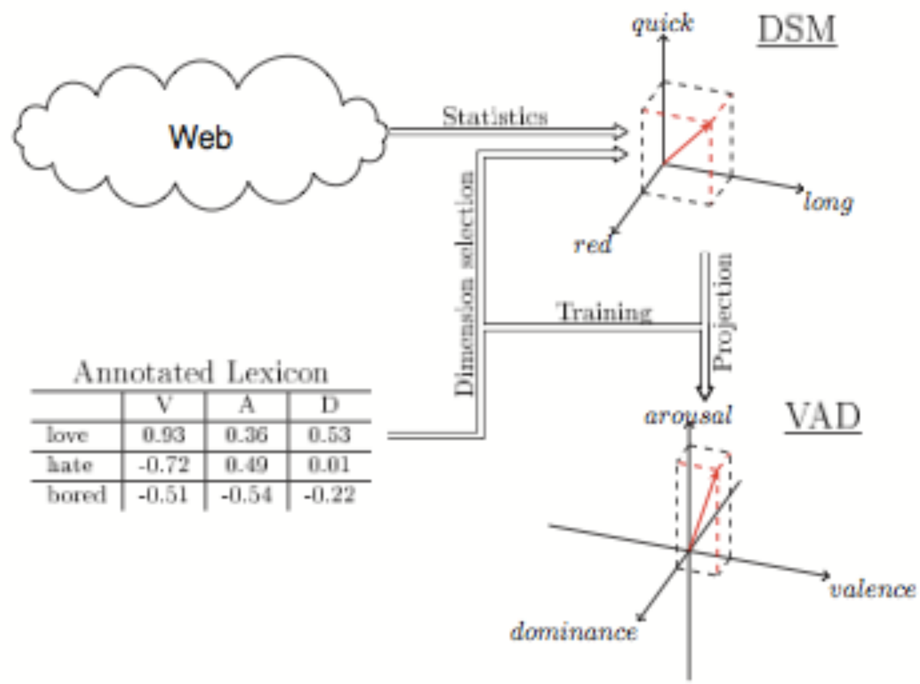
Our lexicon expansion method

Expansion of [Turney and Littman, '02].
Assumption: the valence of a word can be expressed as a linear combination of its semantic similarities to a set of seed words and their valence ratings:

$$\hat{v}(w_j) = a_0 + \sum_{i=1}^{N} a_i \, v(w_i) \, d(w_i, w_j), \tag{1}$$

- w_j : the wanted word
- $w_1 ... w_N$: seed words
- $v(w_i)$: valence rating of word w_i
- a_i : weight assigned to seed w_i
- $d(w_i, w_j)$: measure of semantic similarity between words w_i and w_j

Computations are mappings between layers



Given

- an initial lexicon of K words
- a set of $N < K$ seed words

we can use (1) to create a system of K linear equations with $N+1$ unknown variables:

$$\begin{bmatrix} 1 & d(w_1, w_1)v(w_1) & \cdots & d(w_1, w_N)v(w_N) \\ \vdots & \vdots & \vdots & \vdots \\ 1 & d(w_K, w_1)v(w_1) & \cdots & d(w_K, w_N)v(w_N) \end{bmatrix} \cdot \begin{bmatrix} a_0 \\ a_1 \\ \vdots \\ a_N \end{bmatrix} = \begin{bmatrix} 1 \\ v(w_1) \\ \vdots \\ v(w_K) \end{bmatrix} \quad (2)$$

Solving with Least Mean Squares estimation provides the weights a_i.

Example, $N = 10$ seeds

Order	w_i	$v(w_i)$	a_i	$v(w_i) \times a_i$
1	mutilate	-0.8	0.75	-0.60
2	intimate	0.65	3.74	2.43
3	poison	-0.76	5.15	-3.91
4	bankrupt	-0.75	5.94	-4.46
5	passion	0.76	4.77	3.63
6	misery	-0.77	8.05	-6.20
7	joyful	0.81	6.4	5.18
8	optimism	0.49	7.14	3.50
9	loneliness	-0.85	3.08	-2.62
10	orgasm	0.83	2.16	1.79
-	w_0 *(offset)*	1	0.28	0.28

Sentence Tagging

Simple combinations of word ratings:

- linear (average)

$$v_1(s) = \frac{1}{N}\sum_{i=1}^{N} v(w_i)$$

- weighted average

$$v_2(s) = \frac{1}{\sum_{i=1}^{N} |v(w_i)|}\sum_{i=1}^{N} v(w_i)^2 \cdot \mathrm{sign}(v(w_i))$$

- max

$$v_3(s) = \max_i(|v(w_i)|) \cdot \mathrm{sign}(v(w_z)), \quad z = \arg\max_i(|v(w_i)|)$$

N-gram Affective Models

- Generalize method to n-grams

$$v_i(s) = a_0 + a_1 v_i(unigram) + a_2 v_i(bigram)$$

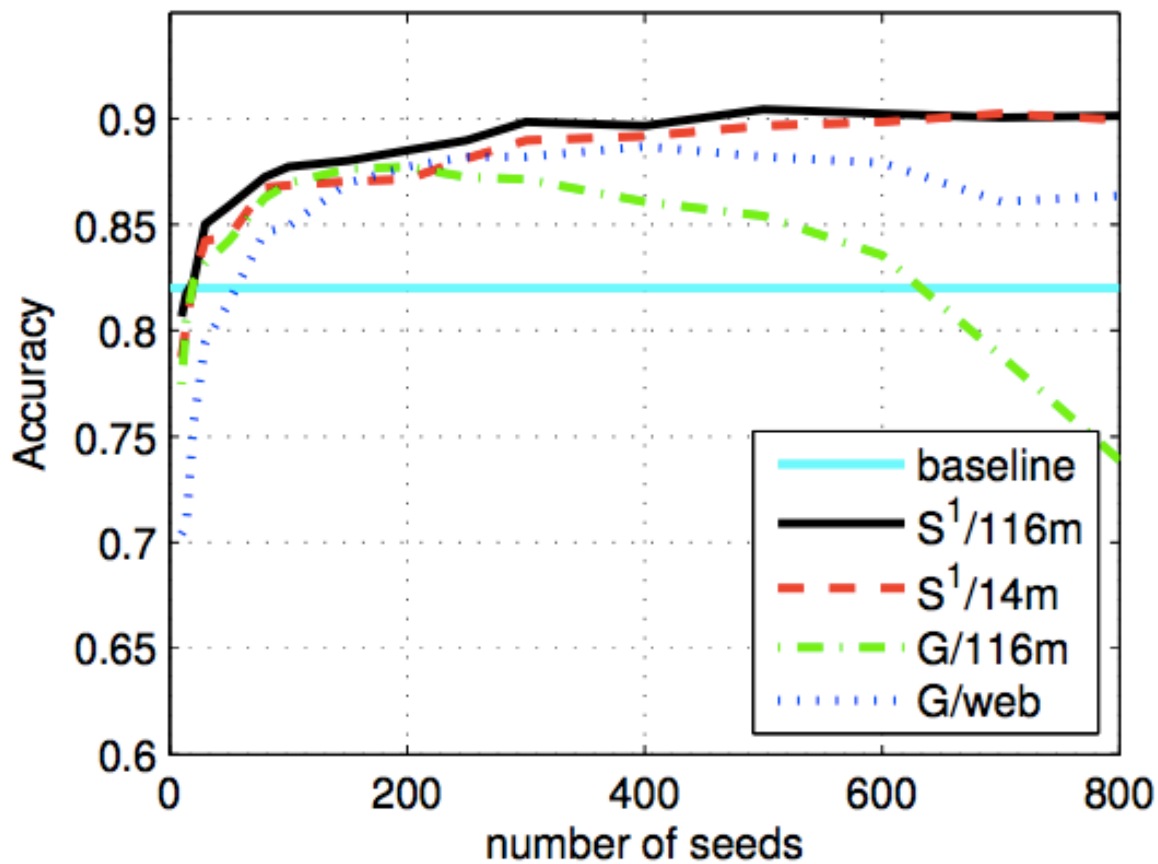
- Starting from all 1-grams and 2-grams, select terms:
  1. Backoff: use overlapping bigrams as default, revert to unigrams based on mutual information-based criterion
  2. Weighted interpolation: use all unigrams and bigrams as default, reject bigrams based on criterion
- In both cases unigrams and bigrams are given linear weights, trained using LMS

Evaluation

- ANEW Word Polarity Detection Task
  - Affective norms for English words (ANEW) corpus
  - 1.034 English words, continuous valence ratings
- General Inquirer Word Polarity Detection
  - General Inquirer words corpus
  - 3.607 English words, binary valence ratings
- BAWLR Word Polarity Detection Task
  - Berlin affective word list reloaded (BAWLR) corpus
  - 2.902 German words, continuous valence ratings
- SemEval 2007 Sentence Polarity Detection
  - SemEval 2007 News Headlines corpus
  - 1.000 English sentences, continuous valence ratings
  - ANEW used for lexicon training
  - 250 sentence development set used for word fusion training

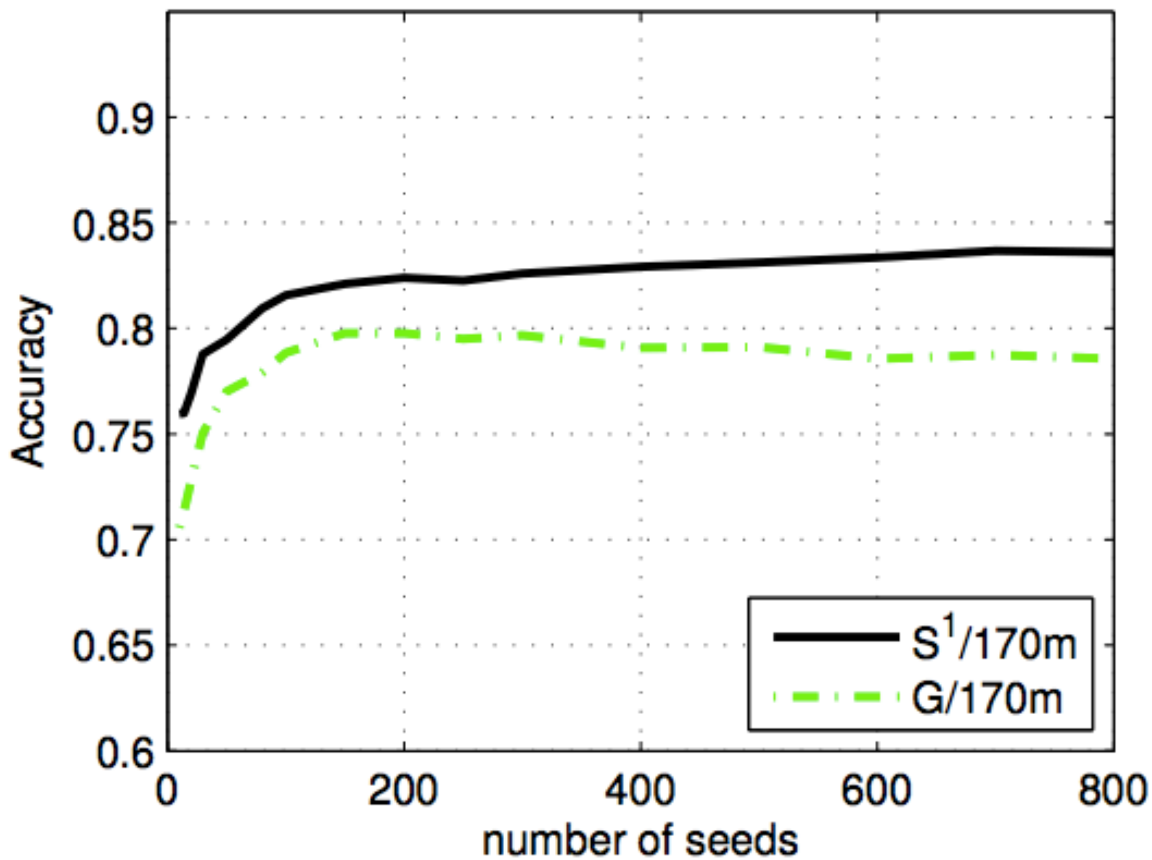
Word Polarity Detection (ANEW)

2-class word classification accuracy (positive vs negative)



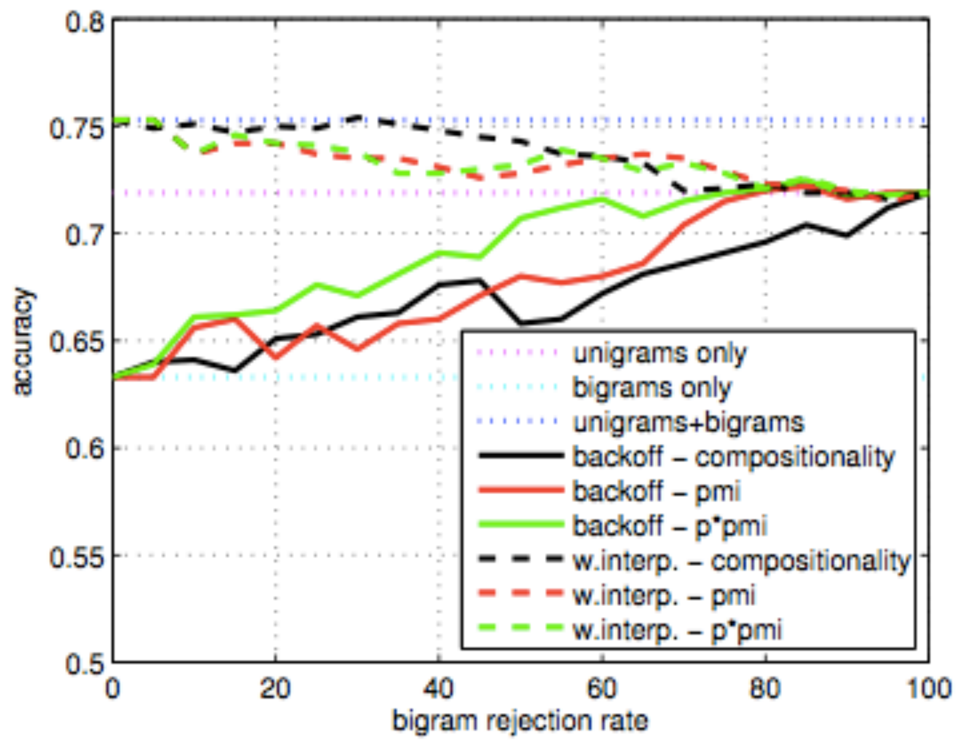
Word Polarity Detection (BAWLR)

2-class word classification accuracy (positive vs negative)



Sentence Polarity Detection (SemEval 2007)

2-class sentence classification accuracy (positive vs negative), vs bigram rejection threshold



ChIMP Sentence Frustration/Politeness Detection

- ChIMP Children Utterances corpus
- 15.585 English sentences, Politeness/Frustration/Neutral ratings
- SoA results, binary accuracy P vs 0 / F vs O:
  - 81% / 62.7% [Yildirim et al, '05]
- 10-fold cross-validation
- ANEW used for training/seeds to create word ratings
- ChiMP words added to ANEW with weight *w*, to adapt to the task
- Similarity metric: Google semantic relatedness
- Only content words taken into account

Politeness: Sentence Classification Accuracy	Fusion scheme		
	avg	w.avg	max
Baseline: P vs O	0.70	0.69	0.54
Adapt $w = 1$: P vs O	0.74	0.70	0.67
Adapt $w = 2$: P vs O	0.77	0.74	0.71
Adapt $w = \infty$: P vs O	**0.84**	0.82	0.75

Frustration: Sentence Classification Accuracy	Fusion scheme		
	avg	w.avg	max
Baseline: F vs O	0.53	0.62	**0.66**
Adapt $w = 1$: F vs O	0.51	0.58	0.57
Adapt $w = 2$: F vs O	0.49	0.53	0.53
Adapt $w = \infty$: F vs O	0.52	0.52	0.52

Summary of Results

- The word-level ratings are very accurate and robust across different corpora
- N-gram sentence-level ratings significantly better than the state-of-the-art, despite the simplistic sentence level fusion model and disregard of syntax/negations
- Adaptation provided good performance on the politeness detection task (linear fusion)
- The baseline model performed best on the frustration detection task (max fusion)

Conclusions

Proposed a high-performing, robust, general-purpose and scalable algorithm for affective lexicon creation

- Investigated linear and non-linear sentence level fusion schemes, showing good but task-dependent performance
- Investigated domain adaptation with good but task-dependent performance (politeness vs frustration detection task)
- Demonstrated that distributional approach can generalize to n-grams

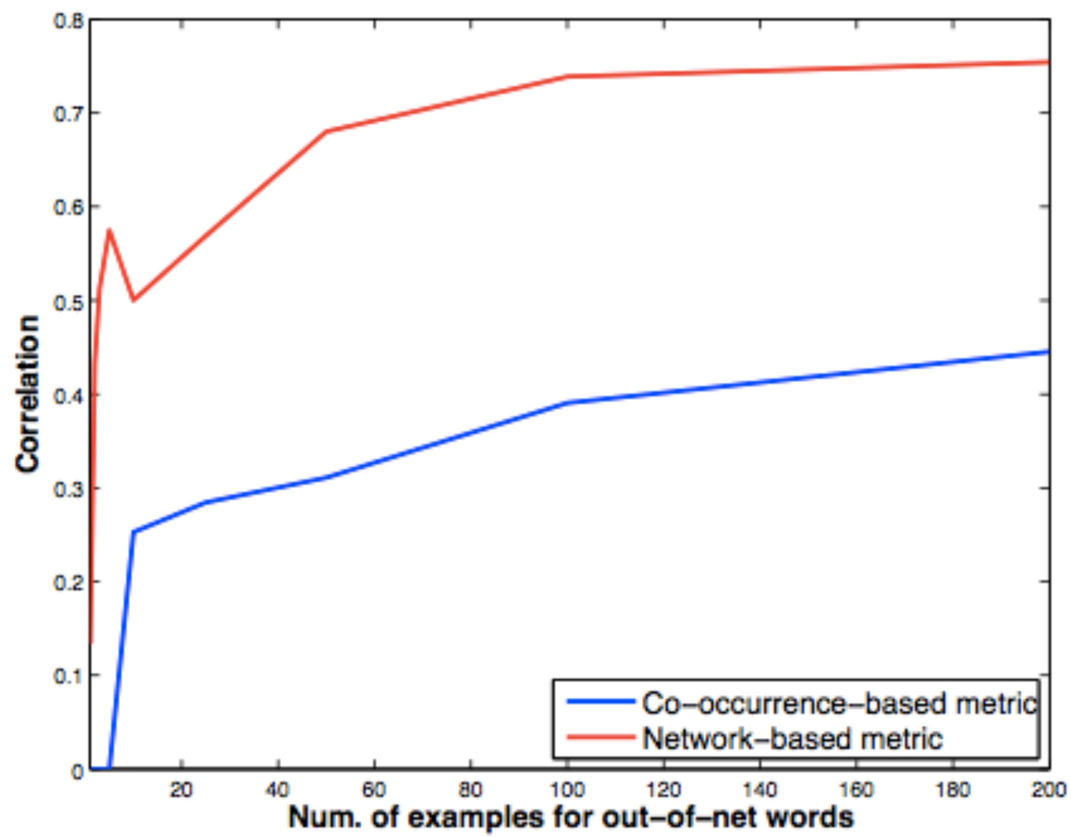
Conclusions

Score Card

Cognitively-motivated semantic models

- Foreground-background classification using attention/ saliency
- Emphasis on induction not classification
- Associations not probabilities/distance
- Mappings between layers
- Hierarchical manifold models not metric spaces
- Multimodal not unimodal

Acquisition of lexical semantics



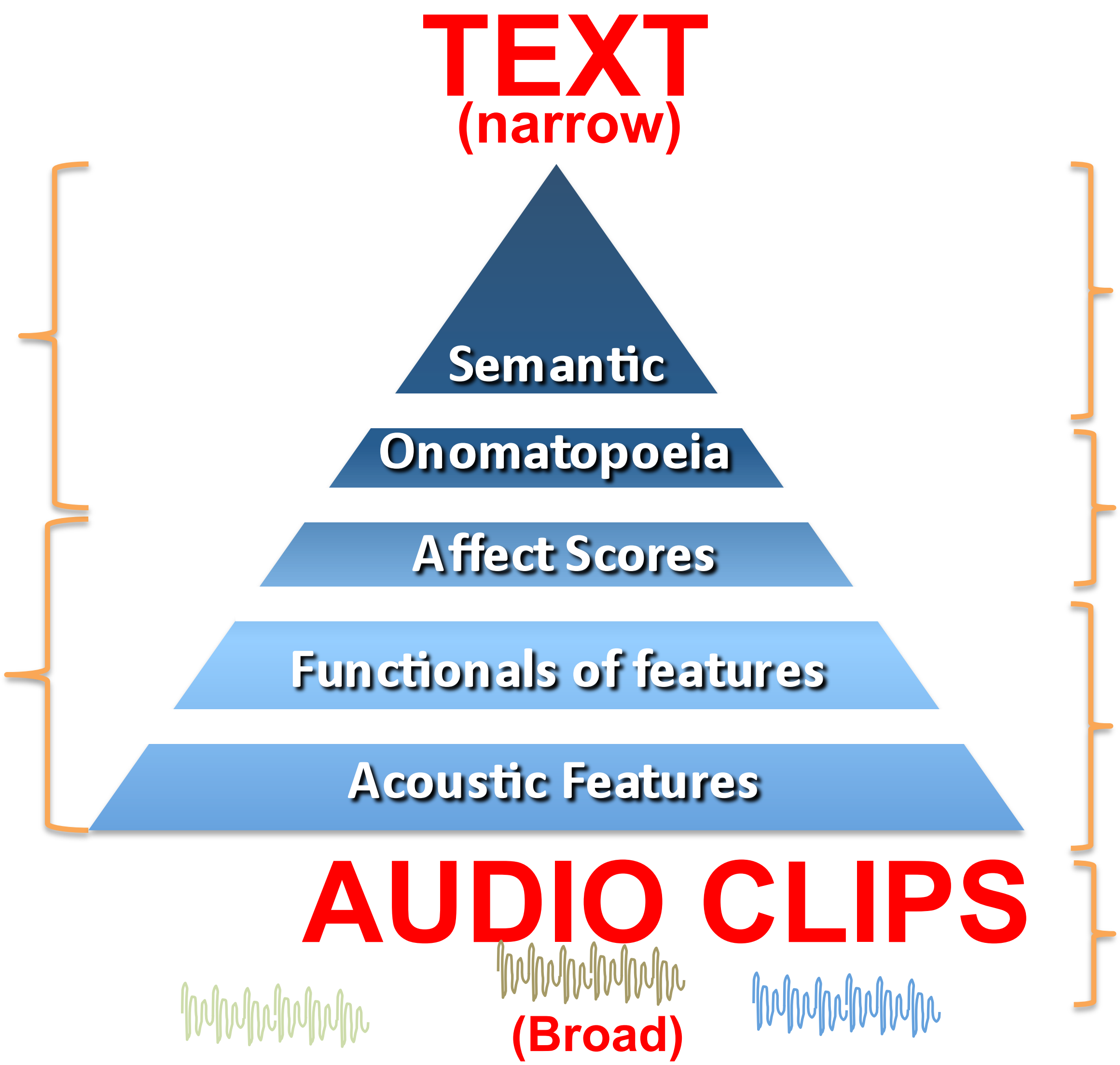
Grand Challenge

Representation Models for Multimedia

- Similarity is the main building block
  - 3 types: similarity w. internal semantic representation, self-similarity over time, similarity in context (biases by world/internal view)
  - Associative network is layer 1 – all computations use this basic representation
- Detectors live in low-dimensional spaces with good geometric properties ("metric")
- Features are labels, labels are features
- Features/labels are organized hierarchically (multiple layers from specific to general, i.e., abstraction)

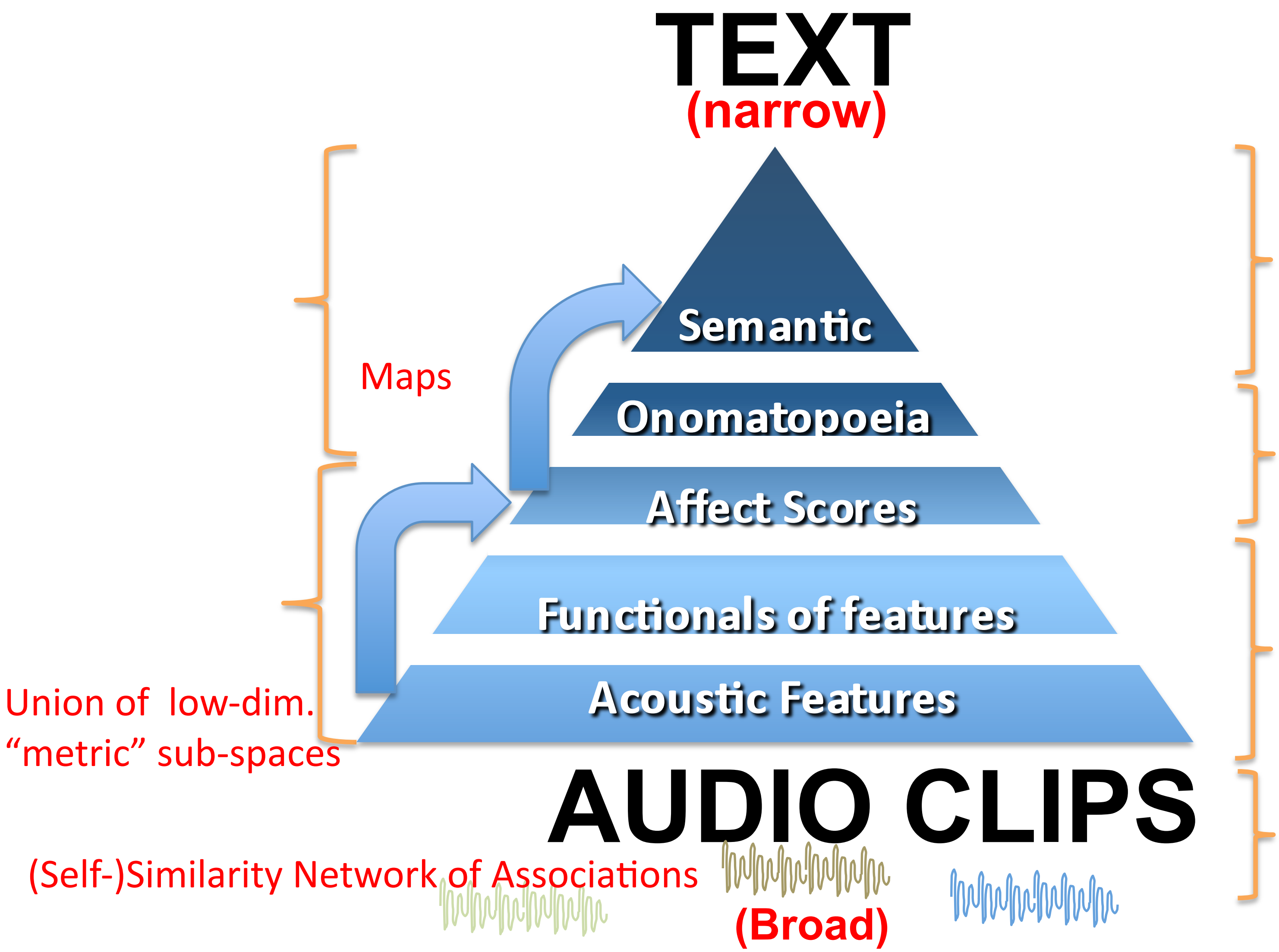
Descriptions of Sounds

[slide by Shiva Sundaram]



Descriptions of Sounds

[original slide by Shiva Sundaram]



Our Timeline

- Unexpectedly good results on semantic similarity tasks using web data
- [E. Iosif, and A. Potamianos, "Unsupervised Semantic Similarity Computation Between Terms Using Web Documents," *IEEE Transactions on Knowledge and Data Engineering*, Nov. 2010]
  - Lucky enough to: 1) work on a semantic similarity task,
    2) directly modeling human cognition
- Goal: reduce web query complexity from quadratic to linear
  [E. Iosif, and A. Potamianos, "Similarity Computation Using Semantic Networks Created From Web-Harvested Data", *Natural Language Engineering*, 2013]
  - Lucky enough not to stop at good initial performance
- Realization:
  - generalization power is in the semantic representation/network
  - multi-tier models: associative network is the 1st tier
- Cognitive science literature [P. Gardenfors, *Conceptual Spaces,* 2000]
  - Low-dimensional "metric" sub-spaces (good geometric properties)
  - Maps and operators defined in this space
- Combine experience from machine learning to come up with a general model