Entropy-Based Evaluation of Decoders

Stefan Steidl

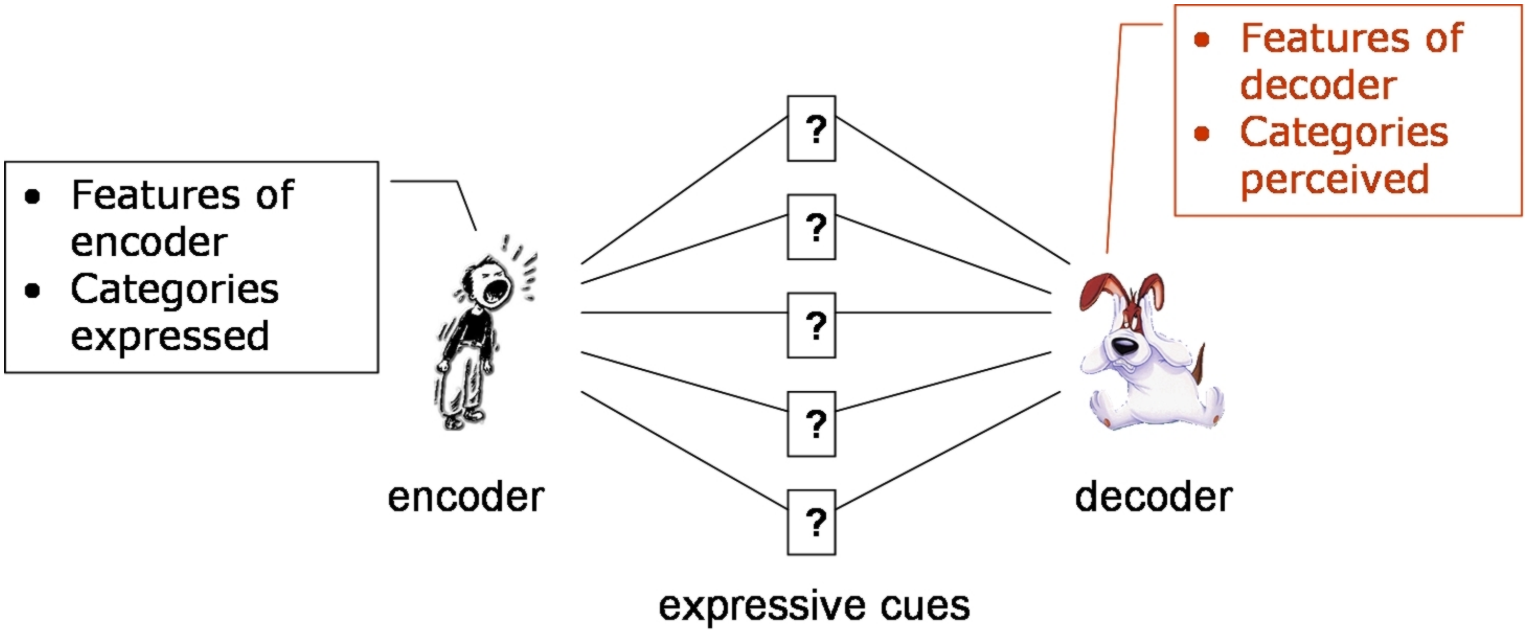
stefan.steidl@informatik.uni-erlangen.de

Humaine WP5 Workshop, Belfast

2nd December 2004

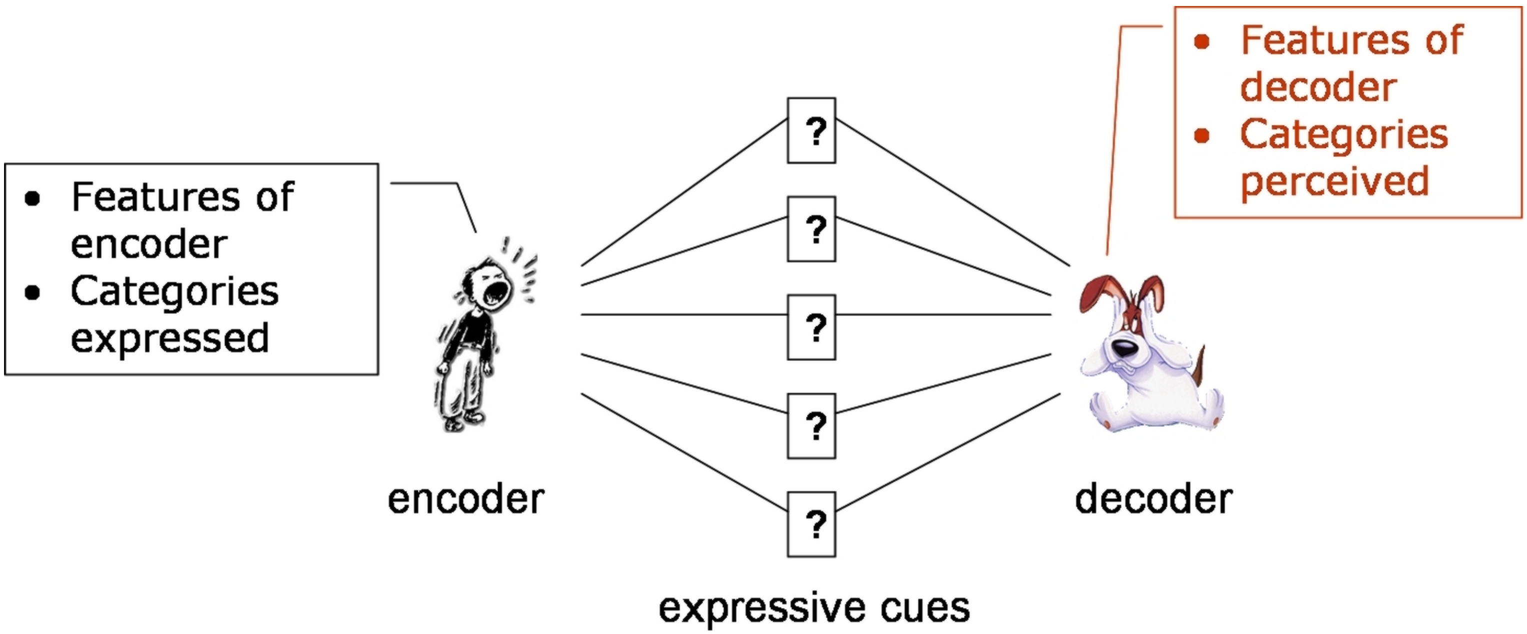
humaine

Emotional Communication Skills



ⓒTanja Bänziger, Geneva Emotion Research Group

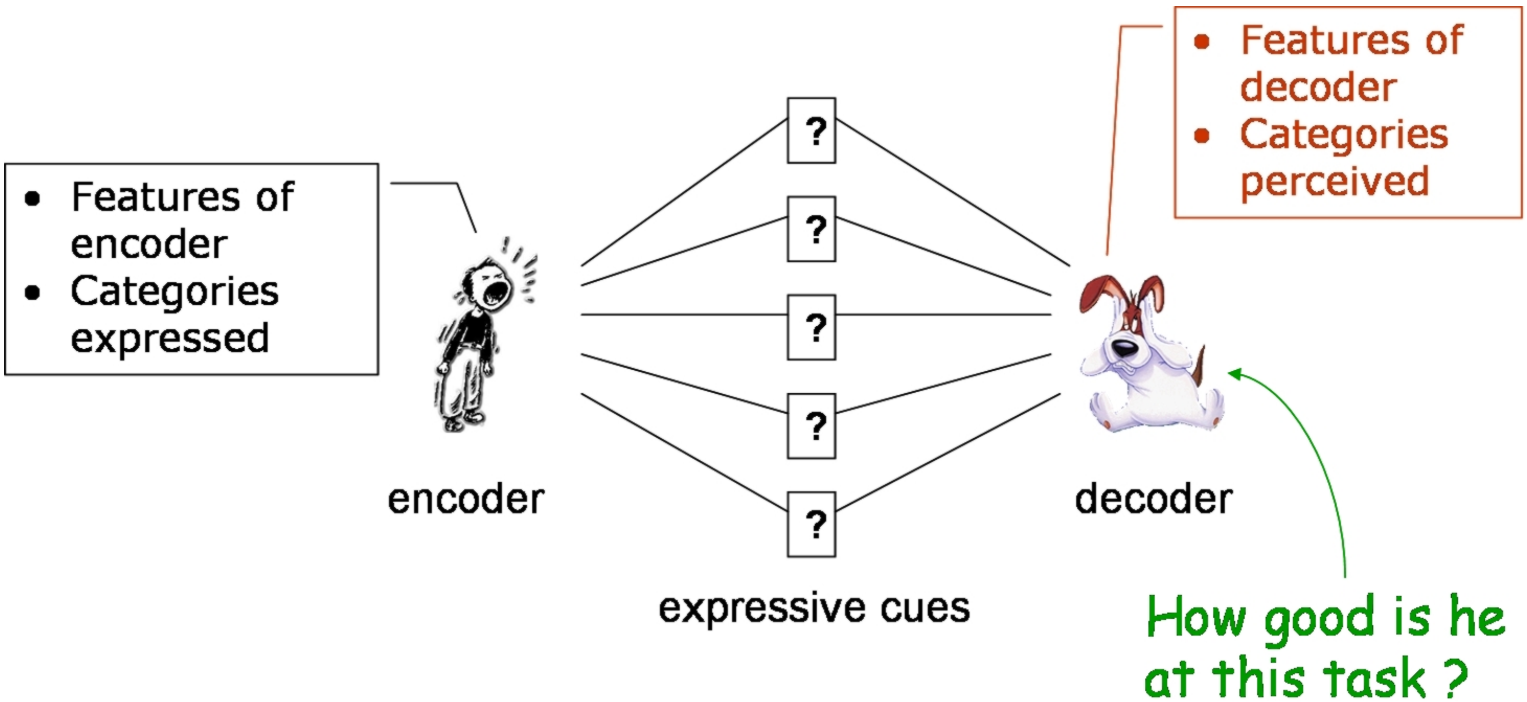
Emotional Communication Skills



©Tanja Bänziger, Geneva Emotion Research Group

- Psychologists: sensitivity of individuals to emotional expressions
- Computer scientists: performance of machine classifiers

Emotional Communication Skills



©Tanja Bänziger, Geneva Emotion Research Group

- Psychologists: sensitivity of individuals to emotional expressions
- Computer scientists: performance of machine classifiers

Evaluation of the Decoder

Traditional:

- (Classwise averaged) recognition rates

Problems:

- (Hard) reference needed
- Dependent on the number and the similarity of the classes
- Confusions of similar emotions are as wrong as confusions of totally different emotions
- Independent of the "quality" of the reference

The Aibo-Emotion-Corpus

- Words labelled by 5 labellers as we do not know which emotion the child expressed
- 4 cover classes: **A**nger, **M**otherese, **E**mphatic, **N**eutral
- Majority voting as reference (3 or more labellers)

emotion	frequency	
anger	1645	3,3 %
5 of 5	191	11,6 %
4 of 5	473	28,7 %
3 of 5	981	59,5 %
emphatic	2528	5,2 %
5 of 5	67	2,6 %
4 of 5	550	21,7 %
3 of 5	1911	75,5 %

emotion	frequency	
motherese	1261	2,6 %
5 of 5	16	1,2 %
4 of 5	510	40,4 %
3 of 5	735	58,2 %
neutral	39182	80,9 %
5 of 5	12654	32,2 %
4 of 5	16112	41,1 %
3 of 5	10416	26,5 %

Entropy-Based Evaluation of the Decoder

- Proposed in our ICASSP 2005 paper
- Creation of soft reference labels $\boldsymbol{l}_{\mathrm{ref}}$

L1	L2	L3	L4
A	A	E	N

$\longrightarrow$

A	M	E	N
0,50	0,00	0,25	0,25

- Calculate the entropy

$$
\begin{aligned}
0 \leq H(\boldsymbol{l}_{\mathrm{ref}}) &= -\sum l_i \cdot \mathrm{ld}\, l_i \\
&= -\left(\frac{1}{2} \cdot \mathrm{ld}\frac{1}{2} + \frac{1}{4} \cdot \mathrm{ld}\frac{1}{4} + \frac{1}{4} \cdot \mathrm{ld}\frac{1}{4}\right) = 1.5 \leq \mathrm{ld}(4) = 2
\end{aligned}
$$

The entropy is a measure for the agreement of the labellers.

Entropy-Based Evaluation of the Decoder

- Decision of the decoder $\boldsymbol{l}_{\mathrm{dec}}$

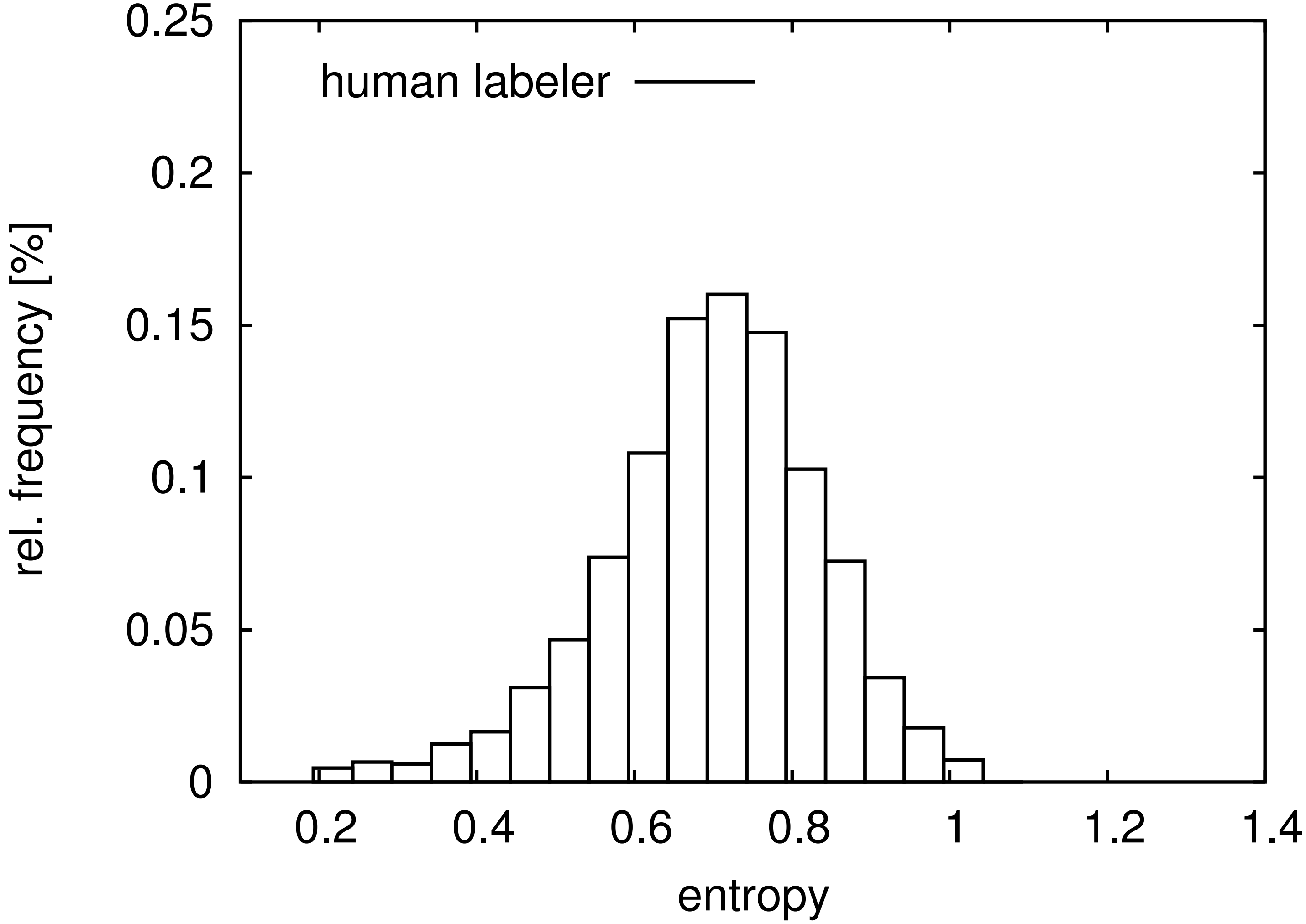
$$\boxed{\begin{array}{c} \text{D} \\ \text{A} \end{array}} \longrightarrow \boxed{\begin{array}{cccc} \text{A} & \text{M} & \text{E} & \text{N} \\ 1{,}00 & 0{,}00 & 0{,}00 & 0{,}00 \end{array}}$$

- 1:1 weighting

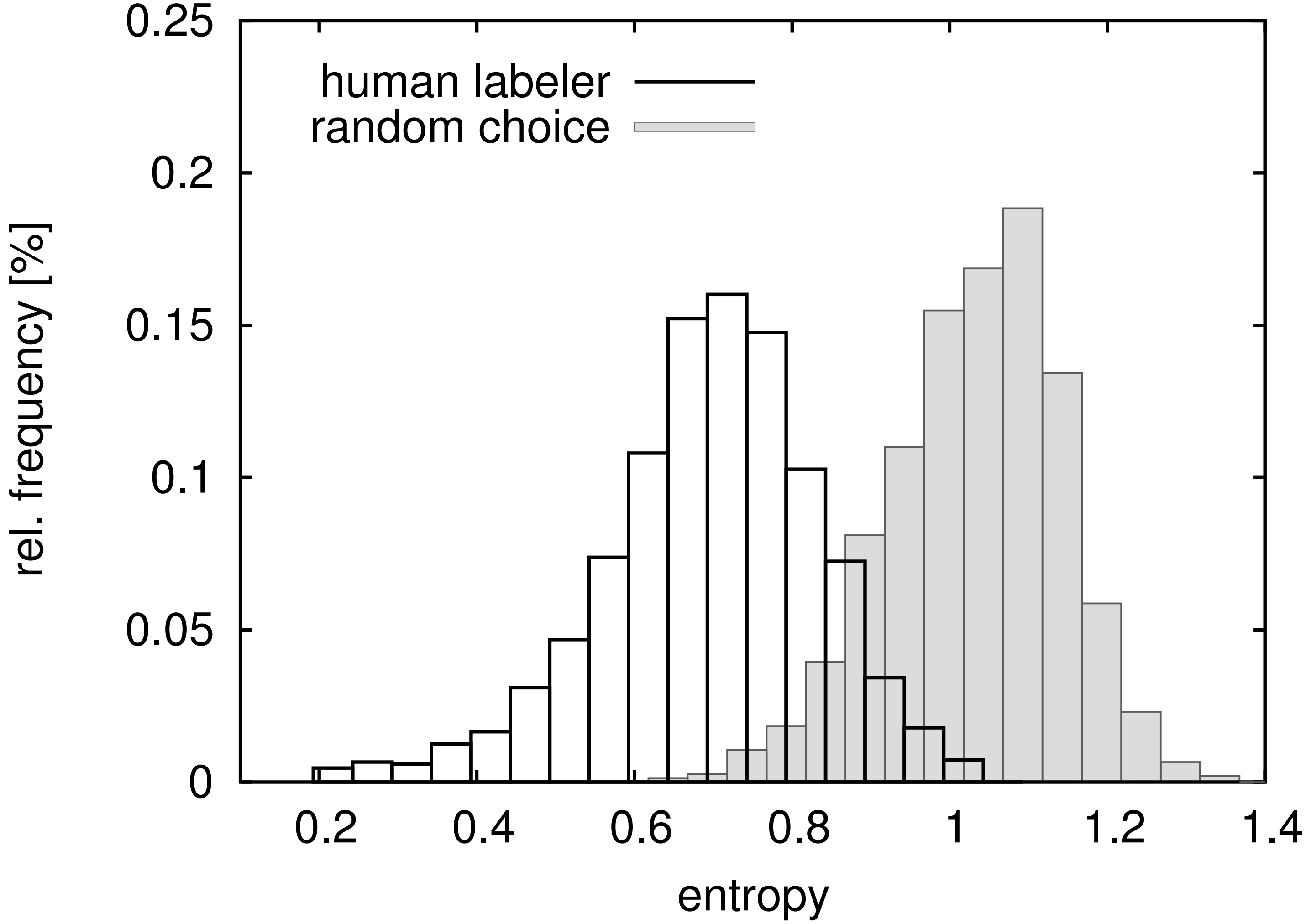
$$\boldsymbol{l} = \frac{1}{2} \cdot \boldsymbol{l}_{\mathrm{ref}} + \frac{1}{2} \cdot \boldsymbol{l}_{\mathrm{dec}}$$

- Calculate the entropy $H(\boldsymbol{l})$
- Calculate the mean entropy of the whole data set or for a certain number of successive samples to plot histograms
- Compare the mean entropy or the histograms for different decoders

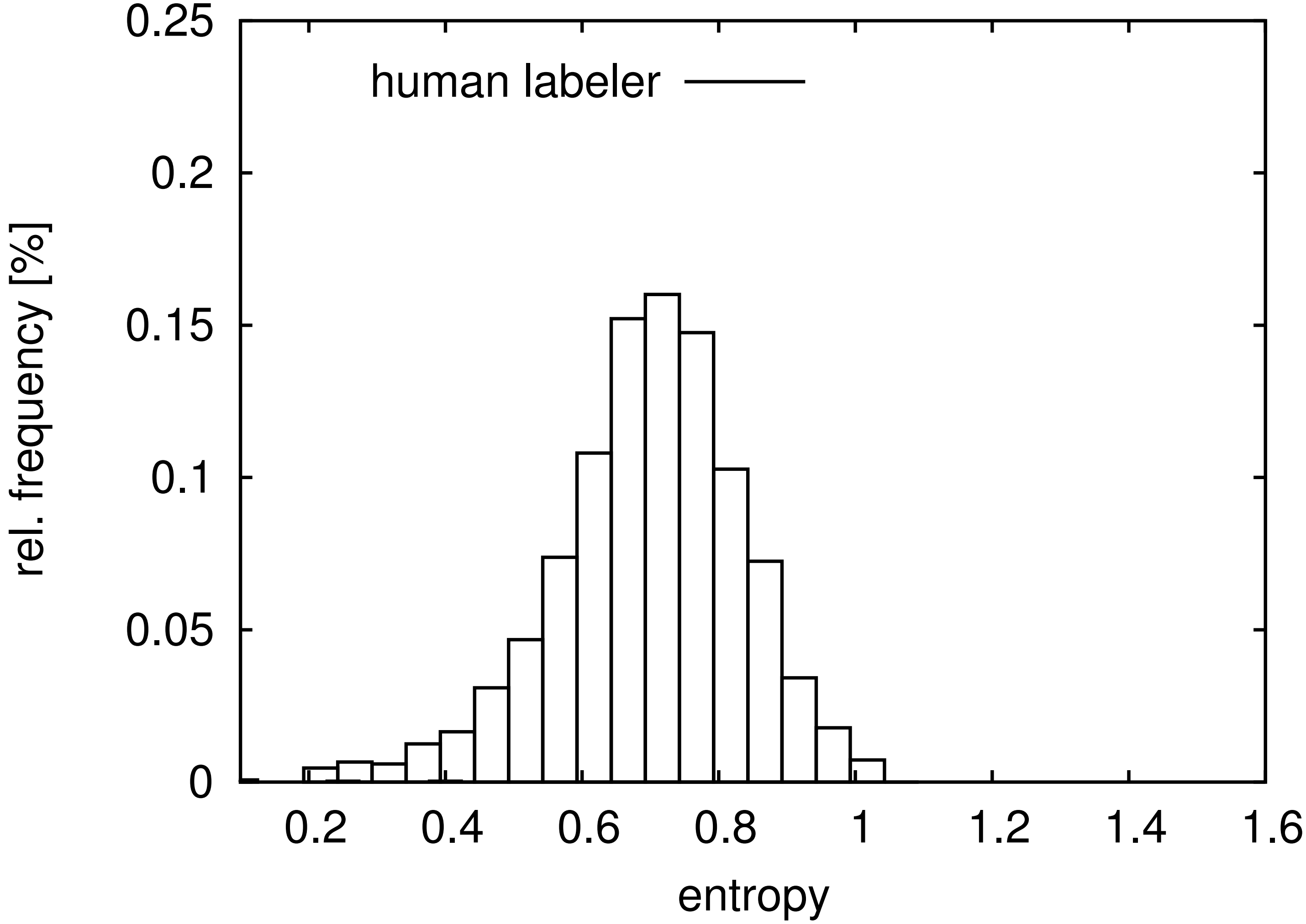
Entropy Histograms



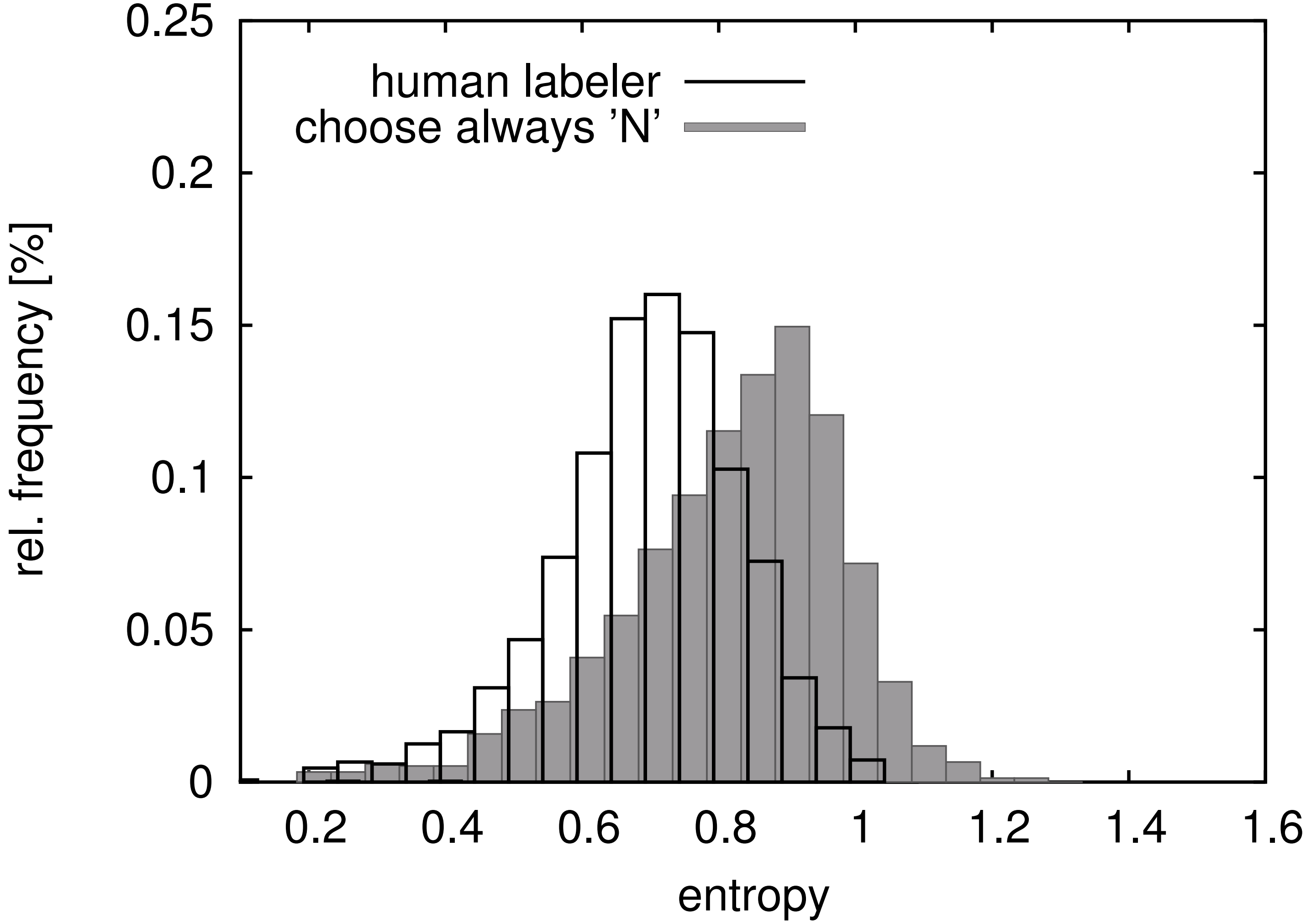
Entropy Histograms



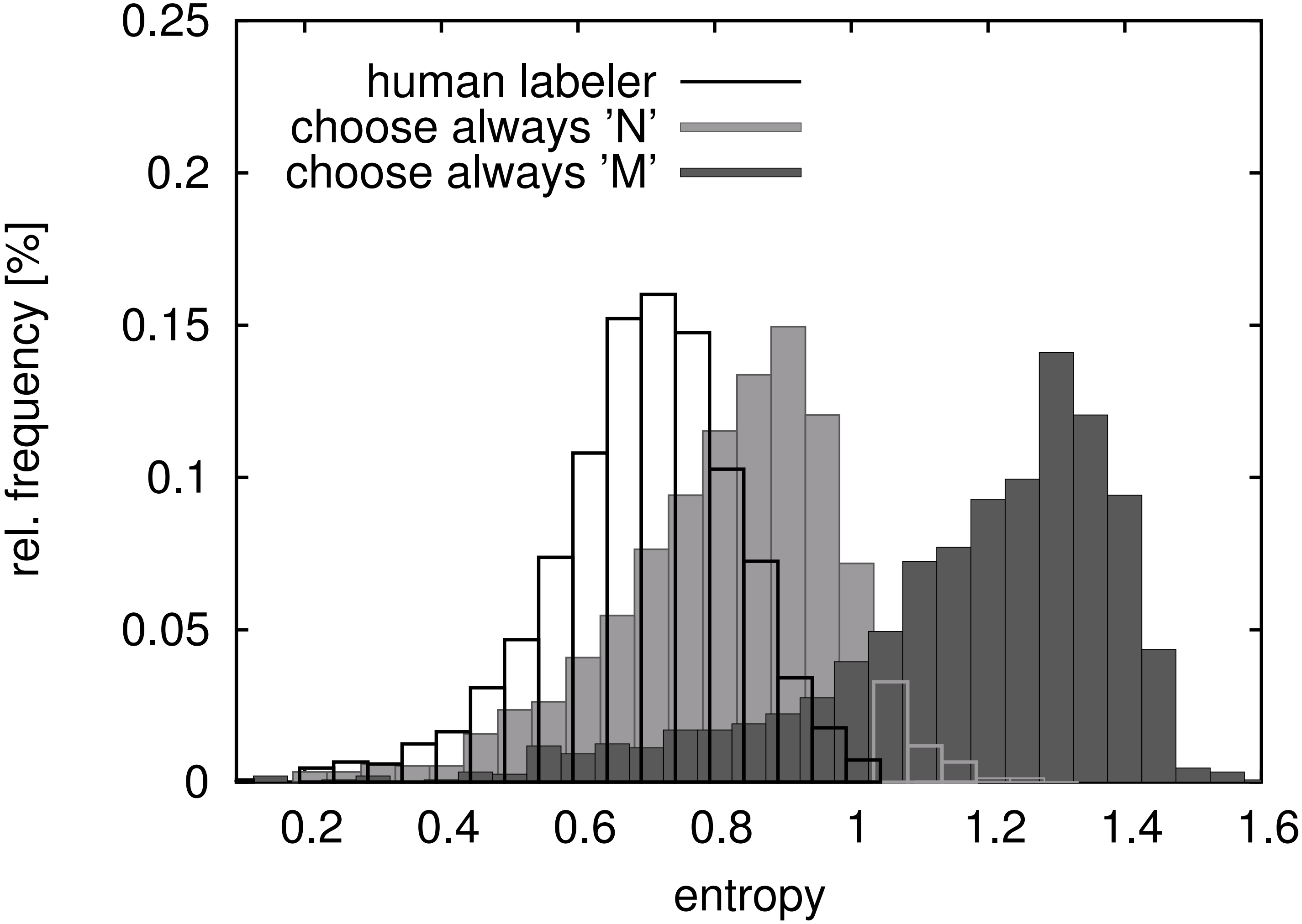
Entropy Histograms



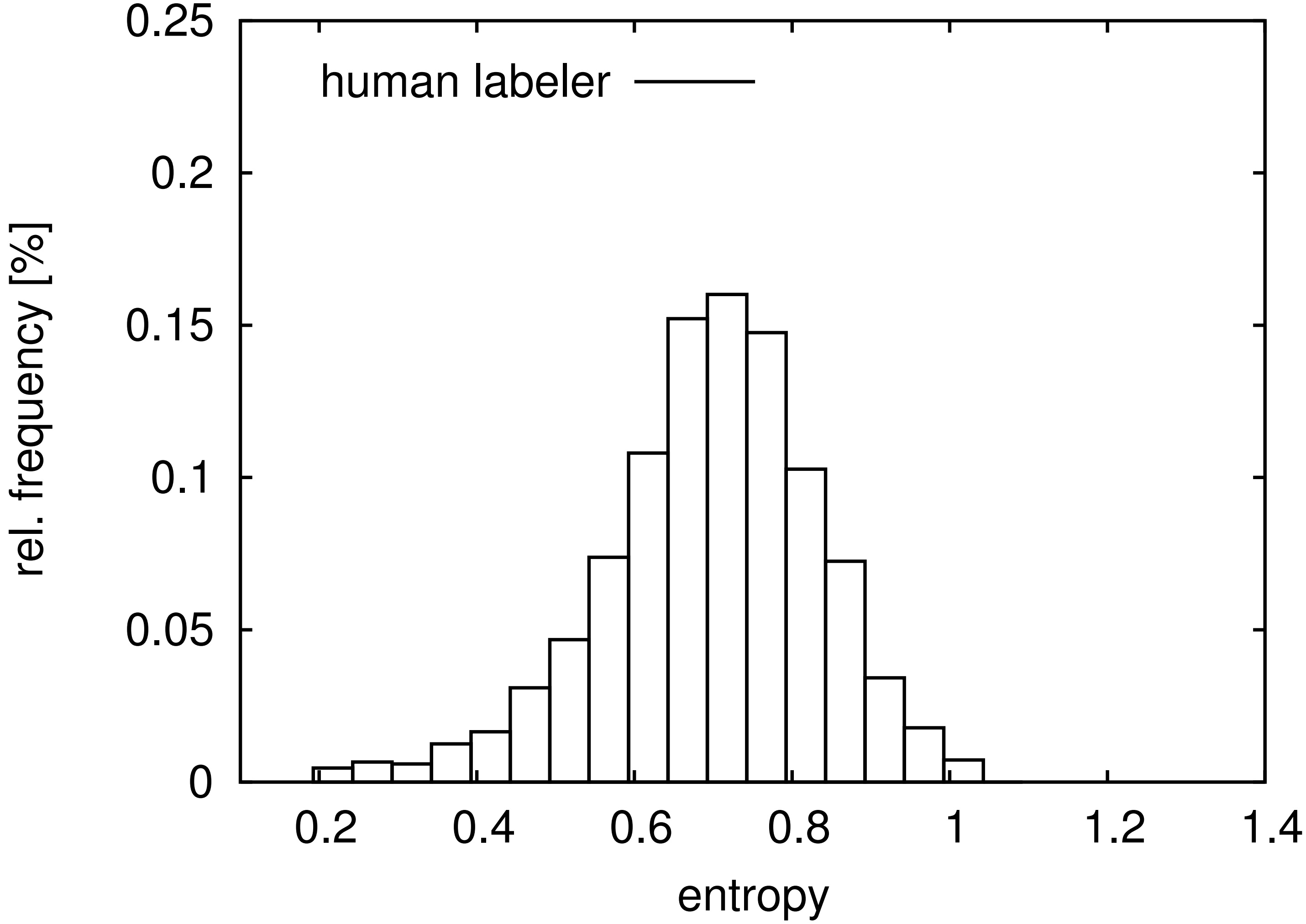
Entropy Histograms



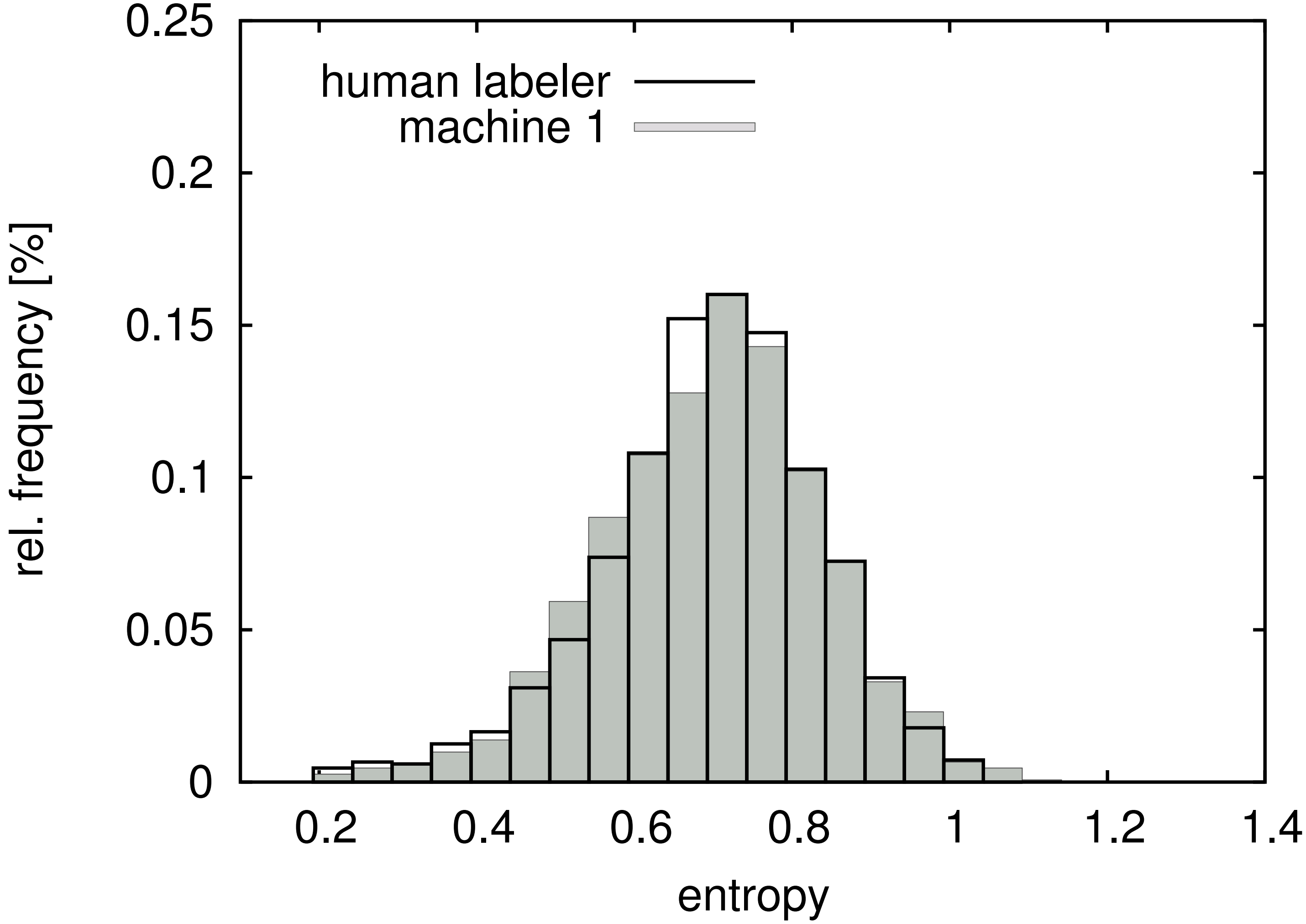
Entropy Histograms



Entropy Histograms



Entropy Histograms



Mean Entropy

decoder	entropy
human labeller	0,721
machine classifier	0,722
all 'neutral'	0,843
all 'emphatic'	1,049
random choice	1,050
all 'anger'	1,127
all 'motherese'	1,196

Mean Entropy

decoder	entropy
human labeller	0,721
machine classifier	0,722
all 'neutral'	0,843
all 'emphatic'	1,049
random choice	1,050
all 'anger'	1,127
all 'motherese'	1,196

Is the recognition problem solved?

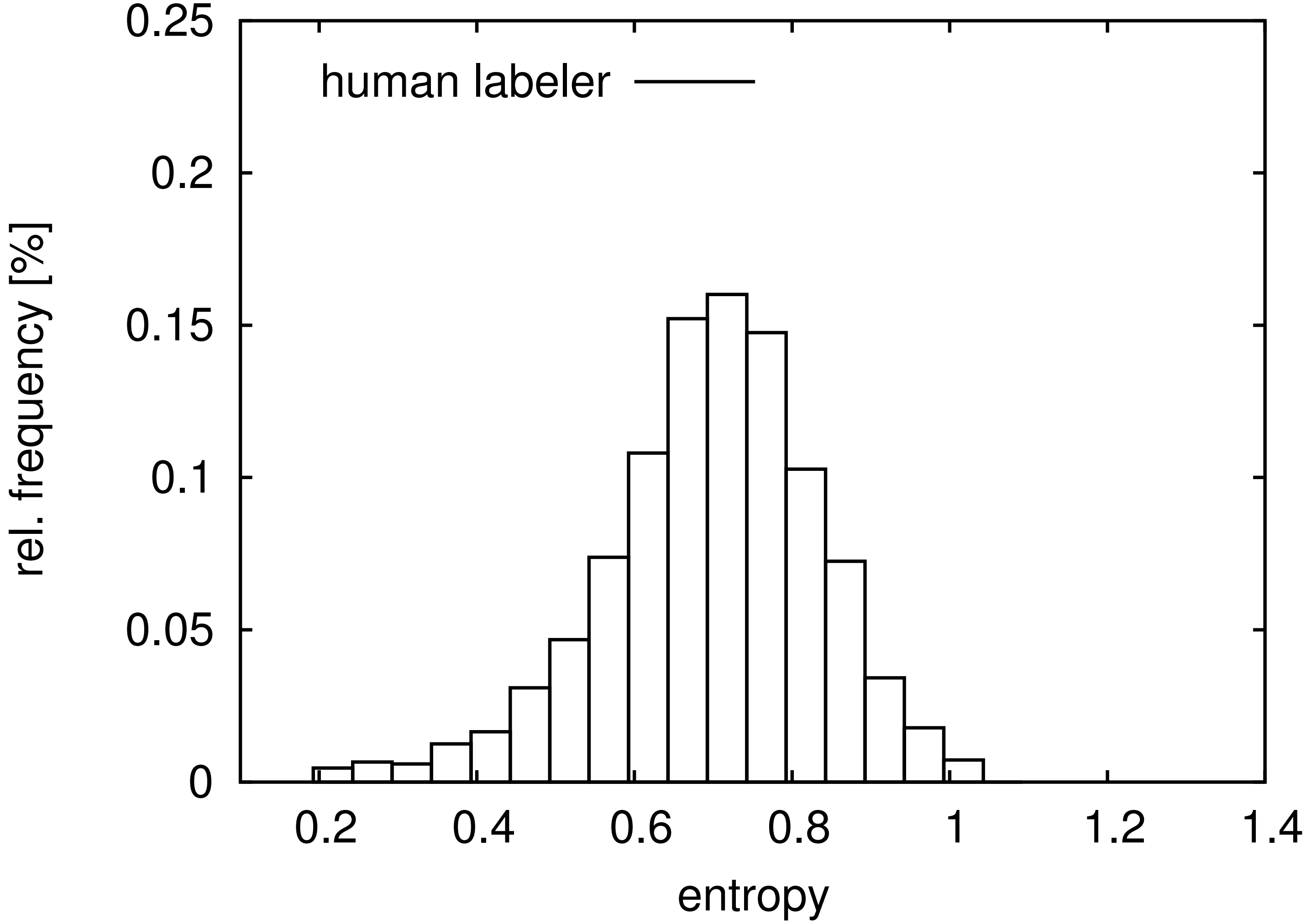
Mean Entropy

decoder	entropy
human labeller	0,721
machine classifier	0,722
all 'neutral'	0,843
all 'emphatic'	1,049
random choice	1,050
all 'anger'	1,127
all 'motherese'	1,196

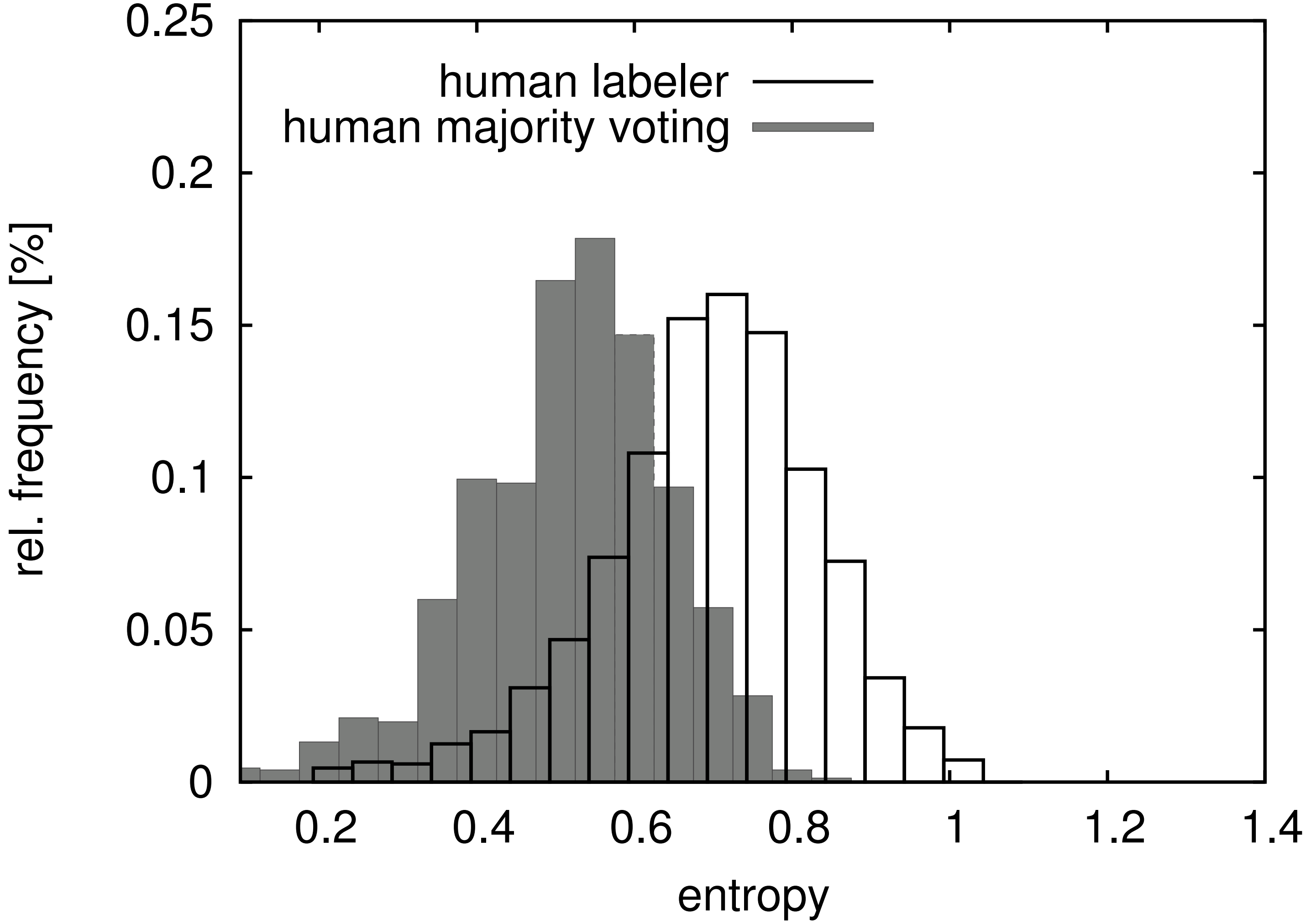
Is the recognition problem solved? Definitely not!

Goal: classification as the majority does

Entropy Histogram



Entropy Histogram



Summary

- Evaluation of decoders with an entropy based measure which
  - ⋆ uses a soft reference
  - ⋆ requires > 2 labellers
  - ⋆ implicitely weights classification "errors" according to the likelihood that both emotions are confused by the labellers

End